

**UNIVERSIDAD COMPLUTENSE DE MADRID**

FACULTAD DE CIENCIAS QUÍMICAS  
Departamento de Bioquímica y Biología Molecular I



**TESIS DOCTORAL**

**Búsqueda racional de nuevas etiquetas de fusión: aplicaciones  
biotecnológicas de los dominios lectina trébol  $\beta$**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Iván Acebrón Ávalos**

Director

José Miguel Mancheño Gómez

**Madrid, 2014**



Departamento de Bioquímica y Biología Molecular I

Facultad de Ciencias Químicas

**Universidad Complutense de Madrid**



**BÚSQUEDA RACIONAL DE NUEVAS ETIQUETAS DE  
FUSIÓN: APLICACIONES BIOTECNOLÓGICAS DE LOS  
DOMINIOS LECTINA TRÉBOL  $\beta$**

TESIS DOCTORAL

Iván Acebrón Ávalos

**Madrid, 2014**



Departamento de Bioquímica y Biología Molecular I

Facultad de Ciencias Químicas

**Universidad Complutense de Madrid**



**BÚSQUEDA RACIONAL DE NUEVAS ETIQUETAS DE  
FUSIÓN: APLICACIONES BIOTECNOLÓGICAS DE LOS  
DOMINIOS LECTINA TRÉBOL  $\beta$**

Memoria presentada por Don Iván Acebrón Ávalos  
para optar al grado de Doctor en Ciencias Químicas

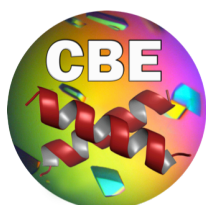
Director del presente trabajo:  
Dr. José Miguel Mancheño Gómez

**Madrid, 2014**



# **BÚSQUEDA RACIONAL DE NUEVAS ETIQUETAS DE FUSIÓN: APLICACIONES BIOTECNOLÓGICAS DE LOS DOMINIOS LECTINA TRÉBOL $\beta$**

Esta memoria, presentada como Tesis Doctoral por Don Iván Acebrón Ávalos para optar al Título de Doctor en Ciencias Químicas por la Universidad Complutense de Madrid, ha sido realizada en el Grupo de Cristalografía y Biología Estructural (CBE) del Instituto de Química Física “Rocasolano” (IQFR) del Consejo Superior de Investigaciones Científicas (CSIC), bajo la dirección del Doctor José Miguel Mancheño Gómez.



**CSIC**  
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



El Dr. José Miguel Mancheño Gómez, científico titular del Departamento de Cristalografía y Biología Estructural del Instituto de Química Física “Rocasolano” (IQFR) del Consejo Superior de Investigaciones Científicas (CSIC), certifica que:

Don **Iván Acebrón Ávalos**, Licenciado en Bioquímica por la Universidad Complutense de Madrid, ha realizado bajo mi dirección el trabajo titulado ***Búsqueda racional de nuevas etiquetas de fusión: aplicaciones biotecnológicas de los dominios lectina trébol  $\beta$***  en el Departamento de Cristalografía y Biología Estructural del Instituto de Química Física “Rocasolano” (IQFR) del Consejo Superior de Investigaciones Científicas (CSIC).

Y, para que así conste a los efectos oportunos, firmo el presente certificado:

En Madrid a            de            de 2014

Dr. José Miguel Mancheño Gómez  
Científico titular del CSIC





Esta tesis fue realizada gracias a la concesión a Don Iván Acebrón Ávalos de una beca predoctoral de Formación de Profesorado Universitario (FPU) (AP2008-03389) durante el periodo comprendido entre el 1 de agosto de 2009 y el 31 de julio de 2013 por parte del Ministerio de Educación, Política Social y Deporte.



*A mis padres y hermanos*



# ***Agradecimientos***

El presente trabajo se ha realizado gracias a la contribución de un gran número de personas. Entre ellas, en primer lugar, querría agradecer profundamente al Dr. José Miguel Mancheño el haberme brindado la oportunidad de dar los primeros pasos en la aventura de la investigación. Juntos hemos puesto nuestro granito de arena en el mundo de las lectinas.

Quisiera agradecer a todos los miembros del Departamento de Cristalografía y Biología Estructural del IQFR su apoyo incondicional durante todos estos años. Especialmente, me gustaría mostrar mi admiración por el Dr. Martín Martínez-Ripoll, ya que su profesionalidad y sabiduría hacen de él un ejemplo a seguir, y por el Dr. Juan A. Hermoso, que me ha dado la oportunidad de seguir formando parte de esta familia un poco más de tiempo. También quiero dar las gracias a la Dra. Beatriz González porque siempre ha estado ahí cuando necesitaba que alguien me escuchase.

No puedo ni podré olvidar a mis compañeros de batalla en el laboratorio durante todos estos años. Todos ellos han contribuido de alguna manera al trabajo que aquí se presenta, dando no sólo apoyo científico, sino moral, mucho más necesario en algunos momentos de esta andadura. Me gustaría agradecer en especial a M<sup>a</sup> Ángela su amistad y todas las experiencias que ambos hemos vivido juntos: congresos, cursos, viajes al sincro, etc. Al resto de las chicas del grupo también las llevaré conmigo siempre porque su ayuda ha sido igualmente incondicional. Por ello, gracias Rocío, Yani, Tere, Ceci, Noe, Elsa, Mer, María, Ainhoa y Juana M<sup>a</sup>. Tampoco puedo dejar de agradecer a Antonio, Javi, Lisandro, Nacho y Sergio su inestimable colaboración. Las anécdotas vividas en los despachos 107 y 117 no se podrían entender sin ellos.

Por otro lado, querría mostrar mi más sincero agradecimiento a las Dras. Rosario Muñoz y Blanca de las Rivas del ICTAN, pues ellas me integraron en su laboratorio como si fuera uno más y me enseñaron a desenvolverse en un laboratorio de Biología Molecular. Además, querría agradecer al resto del equipo (Natalia, Inés, Pura, Laura, Pepe y M<sup>a</sup> Victoria) su cálida acogida. De forma destacada, querría dar las gracias a María (la de Becerril) por todo, ya que ella ha sido uno de los pilares fundamentales durante estos últimos años. Sé que nuestra amistad será duradera y en un futuro compartiremos éxitos científicos.

No quiero dejar de dar las gracias a todas aquellas personas que en algún momento han participado en esta aventura. A la Dra. Margarita Menéndez y a Vicky por enseñarme la

termodinámica de las proteínas. A las Dras. Dolores Solís y M<sup>a</sup> Asunción Campanero y a Ioanna y Lara por “endulzar” esta tesis con su sabiduría en el mundo de los azúcares. Y, por último, a los Dres. Francis Martin y Markus Künzler por surtirnos cuando era necesario de material de trabajo.

Quiero aprovechar esta oportunidad para dar las gracias a muchos de mis profesores (M<sup>a</sup> Luz Álvarez, Ignacio Solá, Isabel de la Mata, Miguel Arroyo, entre otros) que me alentaron con la idea de perseguir mis sueños y adentrarme en el mundillo de la investigación. Y, por supuesto, a mis compañeros de la carrera y del doctorado (María, Isa, Ana, Santi, Nerea, Sarilla, entre tantos otros) porque juntos vamos dando pasos de gigante.

Aunque no pertenecen al mundo académico ni científico, para mí ha sido y es de vital importancia contar con el apoyo de toda mi gente. Muchos ni siquiera saben qué hago exactamente en el laboratorio, pero eso no ha importado para hacerme seguir siempre adelante. Gracias de corazón a todos ellos: Beli, Elena, Lau, Álvaro, Joha, Aída, Baby, Giova, Bev, Fede, Miner y Cris. Muchas gracias, Iván, Miguel, José, Marta, Moni y Jacobo porque en estos últimos años habéis iluminado mi camino.

Finalmente, dejo para el final lo más importante en mi vida, mi familia. Ellos siempre me han apoyado en absolutamente todo y nunca han dudado de mis decisiones. Sé que siempre han querido que alcanzara mis metas, así que espero que esto sirva como prueba de que gracias a ellos una de ellas ya se ha logrado. Esta tesis doctoral va por vosotros. Gracias por todo vuestro cariño, Paco, Pili, Yoli y sobris. Gracias, papá. Gracias, mamá. Os quiero.

*“Mucho aprenderemos en los libros, pero más  
aprenderemos de la contemplación de la Naturaleza,  
causa y ocasión de todos los libros”*

*–Santiago Ramón y Cajal–*

*“The important thing in science is not so much to obtain  
new facts as to discover new ways of thinking about  
them”*

*–Sir William Lawrence Bragg–*





# Índice de contenidos

<b>Resumen</b>	ix
<b>Summary</b>	xi
<b>Capítulo I: Introducción</b>	1
<b>1. Un paseo por el mundo de las lectinas</b>	3
1.1. Breve historia de las lectinas	3
1.2. ¿Dónde se pueden encontrar lectinas?	5
1.2.1. Lectinas de plantas	5
1.2.2. Lectinas de origen animal	6
1.2.3. Lectinas de microorganismos	7
1.3. ¿Qué hacen las lectinas?	8
1.3.1. Lectinas de acción intracelular	9
1.3.2. Lectinas de la superficie celular	10
1.3.3. Lectinas de acción extracelular	11
1.4. El glicocódigo	12
<b>2. ¿Qué nos cuentan las estructuras de las lectinas?</b>	13
2.1. La estructura tipo trébol $\beta$	15
2.2. Evolución del plegamiento trébol $\beta$	17
<b>3. Interacciones lectina-azúcar</b>	19
3.1. Principales interacciones lectina-azúcar	19
3.2. Papel del agua en las interacciones lectina-azúcar	22
3.3. Extensión de los sitios de unión	23
3.4. Multivalencia y cooperatividad	23
3.5. Termodinámica de la interacción lectina-azúcar	24
<b>4. Redescubriendo el reino de los hongos</b>	25
4.1. Diversidad en el reino de los hongos	26

4.1.1. Ascomycota	26
4.1.2. Basidiomycota	27
<b>4.2. Interacción de los hongos con el entorno</b>	27
<b>5. Lectinas fúngicas</b>	28
5.1. Distribución, bioquímica, funciones y estructura	28
5.2. Lectinas fúngicas con estructura trébol $\beta$	31
5.2.1. La lectina LSL de <i>Laetiporus sulphureus</i>	36
5.2.2. La lectina LBL de <i>Laccaria bicolor</i>	39
5.2.3. La lectina SSL de <i>Sclerotinia sclerotiorum</i>	41
5.2.4. La lectina ANL de <i>Aspergillus niger</i>	42
<b>6. Aplicaciones biotecnológicas de lectinas</b>	43
6.1. Estandarización de la producción de proteínas a pequeña y gran escala	44
6.2. Lectinas empleadas como etiquetas de fusión	47
<b>Capítulo II: Objetivos</b>	49
1. Hipótesis de partida y objetivos	51
<b>Capítulo III: Materiales y Métodos</b>	53
1. Visión global	55
2. Expresión y purificación de lectinas fúngicas recombinantes	55
2.1. Obtención de la lectina LSLa de <i>Laetiporus sulphureus</i> y de su correspondiente módulo C-terminal formador de poros	56
2.2. Obtención del módulo lectina N-terminal de la proteína LSLa de <i>Laetiporus sulphureus</i>	58
2.3. Obtención de la lectina LBL de <i>Laccaria bicolor</i> y de su correspondiente módulo lectina N-terminal	60
2.3.1. Producción de la lectina completa de <i>Laccaria bicolor</i>	60
2.3.2. Producción de LBL <sub>152</sub> : el módulo lectina N-terminal de la proteína LBL	62
2.4. Obtención de la lectina SSL de <i>Sclerotinia sclerotiorum</i>	62
2.5. Obtención de la lectina ANL de <i>Aspergillus niger</i>	62

<b>3. Técnicas bioquímicas</b>	64
<b>3.1. Diseño racional de mutantes de LSL<sub>150</sub> mediante mutagénesis dirigida</b>	64
<b>3.2. Análisis de las secuencias de ADN</b>	65
<b>3.3. Espectrometría de masas</b>	65
<b>3.4. Análisis proteómico para la identificación de proteínas</b>	66
<b>3.5. Técnicas electroforéticas</b>	66
3.5.1. Electroforesis en geles de agarosa	66
3.5.2. Electroforesis en geles de poliacrilamida en presencia de dodecilsulfato sódico (SDS-PAGE)	67
<b>3.6. Western blot</b>	67
<b>3.7. Producción de anticuerpos policlonales anti-lectina</b>	68
<b>3.8. Cromatografía analítica de exclusión molecular</b>	69
<b>3.9. Ensayos de actividad enzimática</b>	70
<b>4. Técnicas de caracterización biofísica de proteínas</b>	70
<b>4.1. Espectroscopía de fluorescencia en estado estacionario</b>	70
4.1.1. Espectros de emisión y excitación de fluorescencia	71
4.1.2. Anisotropía de fluorescencia en estado estacionario	72
<b>4.2. Dicroísmo circular</b>	73
<b>4.3. Calorimetría isotérmica de titulación (ITC)</b>	74
<b>4.4. Calorimetría diferencial de barrido (DSC)</b>	76
<b>4.5. Ultracentrifugación analítica</b>	77
<b>5. Análisis de la especificidad de las lectinas</b>	78
<b>5.1. Glycan array screening</b>	78
<b>5.2. Marcaje con biotina</b>	81
<b>5.3. Microarrays</b>	82
<b>6. Estudio estructural de proteínas</b>	84
<b>6.1. Cristalografía de proteínas</b>	85
6.1.1. Cristalización	86
6.1.2. Difracción de rayos X	88
6.1.3. Obtención de datos de difracción y su análisis preliminar	91
6.1.4. Resolución estructural y validación	94

6.1.5. Caracterización estructural de las proteínas de esta tesis	95
<b>6.2. Métodos de predicción de estructura terciaria</b>	97
<b>7. Aplicaciones biotecnológicas de las lectinas con estructura tipo trébol <math>\beta</math></b>	98
7.1. Producción de proteínas recombinantes conteniendo etiquetas de fusión basadas en módulos lectina N-terminales	98
7.2. Estudios de inmovilización sobre agarosa de proteínas fusionadas a LSLt	101
7.3. Microscopía confocal láser de barrido	102
7.4. Caracterización de la unión de lectinas a nanopartículas magnéticas	103
7.4.1. Estudio de la unión de lectinas a nanopartículas magnéticas recubiertas de agarosa (NPA)	103
7.4.2. Orientación y actividad de proteínas inmovilizadas a nanopartículas magnéticas recubiertas de agarosa (NPA)	104
7.4.3. NPA como un método para limpiar un medio de reacción	104
<b>Capítulo IV: Resultados y Discusión. Obtención y caracterización de lectinas fúngicas</b>	105
<b>1. El módulo lectina N-terminal de la proteína LSLa</b>	107
1.1. El módulo lectina N-terminal de la proteína LSLa de <i>Laetiporus sulphureus</i> (LSL <sub>150</sub> ) es una unidad de plegamiento independiente	107
1.2. El módulo LSL <sub>150</sub> actúa como un agente solubilizante natural	108
1.3. El módulo LSL <sub>150</sub> es un trébol $\beta$ que une lactosa	110
1.3.1. Cristalización de LSL <sub>150</sub> y del complejo con lactosa	110
1.3.2. Determinación estructural de LSL <sub>150</sub> y del complejo con lactosa	110
1.3.3. Análisis cristalográfico de la unión a lactosa	113
1.3.4. Comparación del modo de unión del ligando en la familia de los tréboles $\beta$	118
1.3.5. Papel clave de las moléculas de agua en la unión de lactosa	121
1.4. La afinidad de LSL <sub>150</sub> por lactosa en solución es baja	125
<b>2. La lectina LBL de <i>Laccaria bicolor</i></b>	127
2.1. El gen <i>lbl</i> codifica una proteína distinta a la predicha	127

<b>2.2. La lectina LBL se hiperproduce en los cuerpos fructíferos</b>	130
<b>2.3. LBL se proteoliza cuando se produce de forma recombinante</b>	132
2.3.1. Producción de la lectina de <i>L. bicolor</i> en células de <i>E. coli</i> BL21(DE3)	132
2.3.2. Producción de la lectina de <i>L. bicolor</i> en células de <i>E. coli</i> BL21(DE3) que sobreexpresan el complejo GroEL/GroES	134
2.3.3. Producción de la lectina de <i>L. bicolor</i> en <i>Saccharomyces cerevisiae</i>	137
<b>2.4. Producción y purificación del módulo lectina N-terminal de LBL</b>	140
<b>2.5. La estructura de LBL<sub>152</sub> revela a un nuevo trébol <math>\beta</math></b>	141
2.5.1. Cristalización de LBL <sub>152</sub>	141
2.5.2. Determinación estructural de LBL <sub>152</sub> en complejo con lactosa	142
2.5.3. Descripción de la estructura de LBL <sub>152</sub>	145
2.5.4. Estudio detallado del modo de unión de lactosa	147
2.5.5. Las moléculas de agua son parte de la arquitectura del sitio de unión	154
2.5.6. La unión de lactosa induce cambios sutiles en la estructura de LBL <sub>152</sub>	156
2.5.7. La estructura global de LBL <sub>152</sub> no sufre cambios por la unión del ligando	160
<b>2.6. La lactosa ejerce un efecto estabilizante sobre los tréboles <math>\beta</math></b>	162
<b>2.7. Caracterización termodinámica de la interacción de LBL<sub>152</sub> con lactosa</b>	164
<b>3. La lectina SSL de <i>Sclerotinia sclerotiorum</i></b>	167
3.1. La lectina SSL de <i>S. sclerotiorum</i> forma agregados solubles	167
3.2. SSL sería un nuevo trébol $\beta$ con un único sitio de unión a azúcares	169
<b>4. La lectina ANL de <i>Aspergillus niger</i></b>	170
4.1. La lectina ANL de <i>A. niger</i> se obtiene con un rendimiento muy bajo	170
4.2. ANL sería un nuevo miembro de la familia de los tréboles $\beta$	171
<b>Capítulo V: Resultados y Discusión. Estudio de variantes de LSL<sub>150</sub></b>	175
<b>5. Rediseño del sitio <math>\alpha</math></b>	177

<b>5.1. Los mutantes puntuales del sitio <math>\alpha</math> no reconstituyen la actividad lectina</b>	177
5.1.1. Diseño de los mutantes	177
5.1.2. Purificación de los mutantes	179
5.1.3. Cristalización de los mutantes	179
5.1.4. Resolución estructural de los mutantes	180
5.1.5. Análisis estructural del sitio $\alpha$	182
<b>5.2. Los mutantes triples hacen funcional al subdominio <math>\alpha</math></b>	184
5.2.1. Diseño, purificación, cristalización y resolución estructural de los mutantes triples	184
5.2.2. Análisis estructural de la unión de lactosa en el nuevo sitio $\alpha$	187
<b>5.3. Estabilidad y función van en caminos opuestos</b>	192
<b>5.4. El nuevo sitio <math>\alpha</math> tiene una afinidad relativamente baja</b>	197
<b>5.5. Identificación de un nuevo ligando en el sitio <math>\gamma</math></b>	200
<b>6. Rediseño del sitio <math>\beta</math></b>	203
6.1. Diseño, purificación, cristalización y resolución estructural	203
6.2. Efecto de la mutación Ile85His en la unión de lactosa	206
<b>Capítulo VI: Resultados y Discusión. Análisis funcional de la unión a azúcares</b>	207
<b>7. Estudio de la especificidad de LSL<sub>150</sub> y LBL<sub>152</sub></b>	209
<b>7.1. LSL<sub>150</sub> y LBL<sub>152</sub> presentan un perfil de especificidad por azúcares similar</b>	209
7.1.1. Los módulos lectina se producen normalmente con una etiqueta de seis histidinas	209
7.1.2. LSL <sub>150</sub> y LBL <sub>152</sub> tienen predilección por oligosacáridos ricos en poli-LacNAC	215
7.1.3. Las estructuras ricas en LacNAC tienen relevancia biológica	219
7.1.4. Los tréboles $\beta$ fúngicos presentan diferencias en su especificidad	221
<b>7.2. Los complejos con LacNAC revelan nuevas interacciones</b>	222
7.2.1. Cristalización y resolución estructural de los complejos con LacNAC	222
7.2.2. Determinantes estructurales de la unión específica a LacNAC	224

7.2.2.1. LSL <sub>150</sub>	225
7.2.2.2. LBL <sub>152</sub>	226
<b>7.3. Caracterización termodinámica de la unión del disacárido LacNAc</b>	228
7.3.1. LSL <sub>150</sub>	228
7.3.2. LBL <sub>152</sub>	230
<b>8. LSL<sub>150</sub> y LBL<sub>152</sub> podrían reconocer epítomos de insectos y nematodos</b>	232
8.1. Los módulos lectina LSL <sub>150</sub> y LBL <sub>152</sub> son capaces de unir el glicoepítipo LDN	234
8.2. Extractos celulares de <i>C. elegans</i> y <i>D. melanogaster</i> son reconocidos por los módulos lectina	237
<b>Capítulo VII: Resultados y Discusión. Aplicaciones biotecnológicas novedosas</b>	241
<b>9. Producción de proteínas de fusión con dominios trébol <math>\beta</math></b>	243
9.1. Diseño de un protocolo genérico de expresión-purificación basado en el uso de lectinas	244
9.2. LSLt-Bgl: descripción detallada de un ejemplo práctico	247
9.2.1. Antecedentes: la $\beta$ -galactosidasa de <i>Lactobacillus plantarum</i> , una enzima difícil de purificar	247
9.2.2. Resolviendo el problema de la solubilidad de Bgl: ingeniería de proteínas vs proteínas de fusión	252
9.2.3. La formación del oligómero de Bgl no se ve afectada por la etiqueta de fusión	256
9.2.4. Estudio cualitativo de la actividad catalítica de la enzima	258
<b>10. Inmovilización de biomoléculas en matrices de agarosa</b>	259
10.1. Caracterización de LSLt-BTL2 y LSLt-EGFP	261
10.2. Caracterización de la inmovilización de proteínas recombinantes a partículas de agarosa	262
10.3. Efecto de diferentes variables en la unión de LSLt a agarosa	267
10.4. Estabilidad del módulo LSLt bajo diferentes condiciones	268
<b>11. Inmovilización en nanopartículas magnéticas</b>	270

<b>11.1. Las lectinas se unen a nanopartículas recubiertas de agarosa</b>	271
<b>11.2. La unión a la agarosa es direccional y mantiene la actividad de las proteínas inmovilizadas</b>	277
<b>11.3. Las nanopartículas como un sistema de limpieza</b>	280
<b>Capítulo VIII: Conclusiones</b>	283
<b>Bibliografía</b>	289
<b>Anexos</b>	321
<b>Anexo I: Glosario</b>	323
<b>Anexo II: Abreviaturas</b>	331
<b>Anexo III: Publicaciones</b>	335

Nota

Los superíndices del texto hacen referencia a aquellos conceptos cuya definición queda recogida en el glosario



## ***Búsqueda racional de nuevas etiquetas de fusión: aplicaciones biotecnológicas de los dominios lectina trébol $\beta$***

Las lectinas son proteínas con capacidad de reconocer carbohidratos que no poseen un origen inmunológico. Forman un grupo amplio de proteínas que se encuentran ampliamente distribuidas en todos los reinos biológicos. Por esta misma razón, las lectinas están involucradas en gran cantidad de procesos esenciales para las células: comunicación celular, señalización, defensa, patogénesis, etc. Aunque las lectinas procedentes de animales y plantas han sido extensamente estudiadas desde un punto de vista estructural y funcional, el reino de los hongos ha atraído la atención estos últimos años debido a que son una fuente de nuevas lectinas con potenciales aplicaciones en Biomedicina y Biotecnología.

En este contexto, recientemente se caracterizó funcional y estructuralmente una nueva lectina del hongo basidiomiceto *Laetiporus sulphureous*. Esta lectina, denominada LSLa, consta de dos módulos: un módulo lectina N-terminal (LSL<sub>150</sub>) y un módulo C-terminal formador de poros (PFM). El módulo lectina exhibe una estructura tipo trébol  $\beta$  formada por la repetición de tres subdominios ( $\alpha$ ,  $\beta$  y  $\gamma$ ), constituidos cada uno por cuatro hebras  $\beta$ , que generan una estructura globular con pseudosimetría interna de orden tres. La base adquiere la forma de barril  $\beta$ , mientras que la copa está formada por un triplete de horquillas  $\beta$  donde reside la capacidad de unir carbohidratos.

Una búsqueda racional de lectinas homólogas a LSL<sub>150</sub> permitió identificar tres proteínas fúngicas que aún no habían sido caracterizadas: LBL, SSL y ANL. La producción recombinante en células de *E. coli* de estas proteínas sólo fue satisfactoria en la obtención de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub>, que, en definitiva, fueron los que se estudiaron en detalle con el fin de desarrollar aplicaciones biotecnológicas novedosas.

La caracterización de estas lectinas ha incluido la resolución estructural mediante cristalografía de rayos X de los complejos con lactosa o N-acetilactosamina, dos ligandos específicos de estas lectinas, así como un profundo análisis biofísico para valorar la estabilidad y la afinidad de estos sistemas. Además, la especificidad de estas lectinas fue estudiada mediante la moderna tecnología de los *glycan array* masivos, donde se vio que las cadenas de poli-LacNAc son el mejor ligando para estas lectinas. La versatilidad de estos dominios para reconocer ligandos multivalentes demostró que pueden unir extractos celulares de organismos invertebrados, lo que puede dar las primeras pistas acerca de su posible función biológica.

En el caso particular de LSL<sub>150</sub>, se realizó un análisis mediante mutagénesis dirigida para encontrar los determinantes que albergan los sitios de unión para reconocer azúcares. De esta forma, conseguimos restituir la funcionalidad en el subdominio  $\alpha$  que *a priori* era incapaz de unir ninguna molécula de ligando. Este estudio permitió tener una visión global del reconocimiento, que demostró ser un evento de elevada complejidad donde no sólo las interacciones entre el ligando y los aminoácidos del sitio de unión juegan un papel clave, sino que el papel del solvente es un elemento crítico que debe ser considerado.

Finalmente, toda la información recabada nos dio la oportunidad de desarrollar tres aplicaciones biotecnológicas basadas en la utilización de estos dominios como poderosas etiquetas de fusión. En primer lugar, diseñamos un método de expresión y purificación de proteínas recombinantes utilizando cromatografía de afinidad a agarosa. En segundo lugar, empleamos el dominio lectina como una herramienta alternativa para la inmovilización de proteínas de interés industrial. Por último, hemos desarrollado un sistema basado en nanopartículas magnéticas que sirve para anclar proteínas mediante los módulos lectina reversiblemente y de forma orientada, lo que es una ventaja, por ejemplo, en el diseño de biosensores.

## ***Rational Search of New Fusion Tags: Biotechnological Applications of $\beta$ -Trefoil Domains***

Lectins are carbohydrate-binding (glyco)proteins of non-immune origin. They are widespread distributed in nature throughout all biological kingdoms and are involved in many essential cellular and molecular recognition processes, such as cell communication and signaling, defense, host-pathogen interactions, etc. Even though animal and plant lectins have been extensively studied from a functional and structural point of view, Fungi, which include mushrooms, microfungi and yeasts, have attracted broad interest in recent years because they are indeed a promising source for novel lectins with potential biomedical and biotechnological applications.

In this regard, our group has recently characterized functionally and structurally a novel lectin from the basidiomycete fungus *Laetiporus sulphureous*. This lectin, named LSLa, is mainly composed of two modules: an N-terminal lectin module (LSL<sub>150</sub>) and a C-terminal pore-forming module (PFM). The lectin module has a  $\beta$ -trefoil scaffold that consists of a repeat of three subdomains ( $\alpha$ ,  $\beta$  and  $\gamma$ ), each consisting of a four-stranded antiparallel  $\beta$ -sheet. Together, they create a globular structure with internal pseudo-three-fold symmetry that presents a  $\beta$ -barrel capped by a  $\beta$ -hairpin triplet where sugar-binding sites are actually found.

A rational search of LSL<sub>150</sub> homologues allowed us to identify three new fungal proteins with unknown function: LBL, SSL and ANL. The recombinant production in *E. coli* cells only succeeded when we expressed the N-terminal lectin modules LSL<sub>150</sub> and LBL<sub>152</sub> without their accompanying C-terminal modules. Therefore, these two lectins were the only ones that were studied in detail in order to develop new biotechnological applications.

Firstly, the protein structure of these lectins in complex with lactose and N-acetyllactosamine (two specific ligands) was solved by X-ray crystallography at atomic resolution. Secondly, a further biophysical analysis was carried out to evaluate the thermal stability and the sugar affinity of these systems. In addition to this, the sugar specificity was deeply studied using the modern technology of microarrays where hundreds of sugars can be tested simultaneously. The results revealed strong interactions with poly-LacNAc chains, which is important to unravel the biological role of these lectins. Regarding their function, the multivalency these domains show is in fact a key feature to be considered. A microarray

where some model glycoproteins and cell extracts from invertebrate organisms were immobilized gave also some clues about it.

In order to find the sugar-binding determinants of LSL<sub>150</sub>, we performed a site-directed mutagenesis analysis upon the  $\alpha$ -subdomain, in which no sugar molecule have been bound *in vitro*. The generation of the sugar-binding capacity in the  $\alpha$ -subdomain provided us with a complex vision of the sugar recognition process. Not only direct contacts between the ligand and the protein residues are important, but also the solvent plays a crucial role in sugar binding. In fact, the results indicate that some water molecules behave as an extension of the sugar-binding sites.

Finally, with all the collected information, we have developed three innovative biotechnological applications based on the use of these two modules as powerful fusion tags. In this sense, we designed an expression-purification methodology for recombinant proteins using a single-step agarose-affinity chromatography. Subsequently, the  $\beta$ -trefoil lectin domain was used as an alternative tool to immobilize proteins with industrial interest. In the end, we have put in practice a novel system based on the combination of agarose-coated magnetic nanoparticles and  $\beta$ -trefoil lectin domains. As expected, the immobilized target proteins are well oriented onto the nanoparticle surface and retained their activity, which, for instance, is considered an advantage for biosensor designing.

# CAPÍTULO 1



## INTRODUCCIÓN



## I

## INTRODUCCIÓN

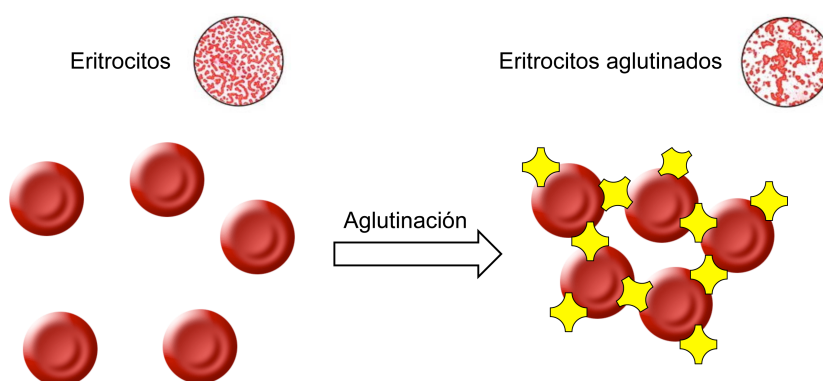
## 1. Un paseo por el mundo de las lectinas

### 1.1 Breve historia de las lectinas

Las lectinas<sup>22</sup> comprenden un grupo de proteínas estructuralmente diverso y con multitud de funciones biológicas asignadas (Rini, 1995). Por esta razón, llegar a un consenso para establecer una definición del término lectina no fue tarea sencilla. La definición que propusieron Peumans y van Damme para las lectinas de plantas se acerca bastante a la descripción que se maneja actualmente: *"In our opinion the presence of at least one noncatalytic domain that binds reversibly to a specific carbohydrate is the only prerequisite for a protein to be named a lectin. Consequently, plant lectins can be defined as all plant proteins that possess at least one noncatalytic domain that binds reversibly to a specific mono- or oligosaccharide."* [En nuestra opinión, la presencia de al menos un dominio no catalítico que se una reversiblemente a un azúcar específico es el único prerequisite para que una proteína sea considerada una lectina. En consecuencia, las lectinas de plantas pueden definirse como todas aquellas proteínas de plantas que posean por lo menos un dominio no catalítico que una reversible y específicamente mono u oligosacáridos.] (Peumans y van Damme, 1995). Para completar esta definición es necesario indicar que estas proteínas no poseen un origen inmunológico, lo que fue parte de la controversia que existió en los primeros estudios que se llevaron a cabo sobre el estudio de lectinas, iniciados a finales del siglo XIX.

La lectinología, como así se ha venido a llamar a la ciencia que trata sobre el estudio de las lectinas, se inició con ensayos cuyo objeto de estudio era principalmente la sangre. Fue en el año 1860 cuando el médico Mitchell observó por primera vez cómo el veneno de la serpiente de cascabel (*Crotalus durissus*) producía la coagulación de sangre procedente de paloma (Mitchell, 1860). Más adelante, se dio cuenta de que aquel fenómeno no se debía a la actuación de ningún agente de coagulación, sino a un proceso de aglutinación<sup>2</sup> celular por efecto de algún componente del veneno (**Figura 1.1**). Al poco tiempo, en 1888, Stillmark volvió a observar ese mismo fenómeno con un extracto de la semilla del ricino (*Ri-*

*cinus communis*). Stillmark dio un paso más allá y propuso en su tesis doctoral que esta sustancia con actividad aglutinante de eritrocitos se trataba de una proteína tóxica que denominó ricina (Stillmark, 1888). A raíz de un mayor número de experimentos en los que, empleándose este tipo de proteínas procedentes tanto de animales como de plantas, se obtenían resultados similares a los anteriores sobre distintas muestras de sangre, se hizo popular el término hemaglutinina para definir a este determinado grupo de proteínas. Estas hemaglutininas presentaban la capacidad de actuar como agentes precipitantes, eran capaces de seleccionar distintos tipos de eritrocitos y se inhibían en presencia de pequeñas moléculas (haptenos) de diferente complejidad química y de origen glucídico (carbohidratos o azúcares<sup>7</sup>). Todas estas características llevaron a pensar que estas proteínas podrían ser consideradas como un tipo especial de anticuerpos.



**Figura 1.1.** Reacción de aglutinación de eritrocitos mediada por lectinas (amarillo). En la parte superior se indica el aspecto de una muestra de glóbulos rojos en buen estado (izquierda) y tras la reacción de hemaglutinación (derecha).

En 1954, sin embargo, Boyd introdujo un nuevo nombre para designar a estas proteínas (Cumings y Pierce, 2014), ya que, de acuerdo a su criterio, la capacidad que presentan para unir específicamente un grupo determinado de moléculas no debía ser considerada condición suficiente para definir las como anticuerpos. Los anticuerpos, a diferencia de estas hemaglutininas, son producidos por el sistema inmunológico con la finalidad de hacer frente a un proceso infeccioso. Por esta razón, Boyd propuso el término lectina, que deriva del latín *legere*, que significa literalmente escoger o seleccionar. Así se dejaba constancia de la capacidad que presentan estas proteínas para seleccionar un tipo específico de moléculas, concretamente carbohidratos, y se desvinculaban por completo de las proteínas involucradas en la respuesta inmunológica. Hoy en día, las lectinas se han establecido como un grupo bien definido, diferenciándose claramente de (1) las inmunoglobulinas específicas de carbohidratos, (2) las enzimas que usan carbohidratos como sustratos y (3) las pro-



teínas transportadoras de mono u oligosacáridos.

A partir de finales de los años sesenta, el estudio funcional y estructural de las lectinas creció aceleradamente. En 1972 se publicó la primera estructura tridimensional de una lectina: la concanavalina A (ConA) de la planta *Canavalia ensiformis* (Rutenber *et al.*, 1991; Drickamer, 1997). Posteriormente, en 1989 llegó la primera estructura de un complejo de una lectina unida a una molécula de azúcar, concretamente, la ConA unida a metil- $\alpha$ -D-manopiranosido (Derewenda *et al.*, 1989), y en 1993, la primera estructura de una lectina animal, una galectina<sup>16</sup> humana unida a lactosa (Mattos, 2002). El conocimiento acumulado durante todos estos años ha permitido proponer una clasificación de las lectinas en diferentes familias con criterios tanto estructurales como funcionales, ordenando así el complejo mundo de la lectinología.

## 1.2 ¿Dónde se pueden encontrar lectinas?

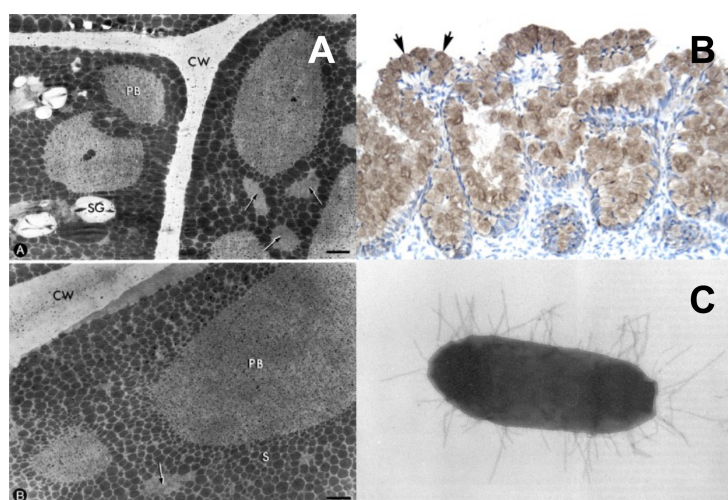
Teniendo en cuenta que el reconocimiento de azúcares es un proceso que ocurre en contextos biológicos muy variados, no es de extrañar que las lectinas aparezcan en todos los reinos de la naturaleza, desde bacterias y arqueas a organismos eucariotas superiores. Su distribución en tejidos u órganos es variable y dependiente de distintos factores como, por ejemplo, la etapa de desarrollo, la edad o condiciones patológicas determinadas. A continuación, se va a describir brevemente la distribución de lectinas en los diferentes reinos biológicos.

### 1.2.1. Lectinas de plantas

Las primeras lectinas objeto de estudio procedían en su mayoría del reino vegetal. Posiblemente, la rápida y fácil disponibilidad de material biológico contribuyó a ello. Es fácil entender, por tanto, que la primera lectina que se describió fuera la ricina, procedente de la semilla del ricino (*R. communis*), y la primera cuya estructura tridimensional fue determinada, la ConA del frijol (*C. ensiformis*). Se ha descrito una gran variedad de lectinas de numerosas especies de plantas, siendo más de la mitad de la familia de las leguminosas (Petukhov *et al.*, 1999). La mayoría de las plantas acumula lectinas en las semillas (Cavada *et al.*, 1998), además de en otros órganos con fines de almacenamiento como, por ejemplo, en la corteza, raíces o bulbos. Normalmente, estas lectinas son citosólicas y se encuentran asociadas a otras proteínas en los llamados cuerpos proteicos<sup>10</sup> (Rüdiger, 1998) (**Figura 1.2A**). La cantidad de lectinas presentes en plantas es muy variable (Rüdiger y Gabius,

2001), aunque siempre mayor que la encontrada en otro tipo de organismos, lo que ha supuesto una ventaja para la caracterización estructural, en la que se requiere mayor cantidad de material de partida. En último término, cabe resaltar que en algunas plantas, sobre todo en las leguminosas, se han identificado mezclas de isoformas de lectinas, conocidas como isolectinas. Éstas suelen tener la misma estructura tridimensional, pero difieren en la especificidad o en algunas propiedades fisicoquímicas como la carga eléctrica neta (Sharon y Lis, 2007).

Las lectinas fúngicas solían incluirse históricamente en el grupo de las lectinas de plantas debido a la estrecha relación que ha existido siempre entre plantas y hongos. Sin embargo, desde que en 1989 se purificasen las primeras lectinas de los cuerpos fructíferos de los hongos *Agaricus campestris* y *Agaricus bisporus* (Goldstein y Poretz, 1986), las lectinas fúngicas han ido conformando un grupo independiente por méritos propios. Las lectinas de hongos serán tratadas en detalle más adelante en el apartado 1.5 de la *Introducción*.



**Figura 1.2.** Diferentes fuentes biológicas de lectinas. **A.** Sección fina del cotiledón de la planta *Glycine max* (soja) donde se muestra la localización de la lectina SBA mediante marcaje con anticuerpos anti-SBA y revelado con partículas de oro (*puntos negros*). Las flechas indican los cuerpos proteicos (PB) no marcados. **B.** Visualización inmunohistoquímica del receptor de asialoglicoproteínas del intestino de conejo. **C.** Micrografía de *Escherichia coli* en la que se observan las fimbrias donde se encuentran localizadas las lectinas.

### 1.2.2. Lectinas de origen animal

Las lectinas animales, incluyendo las humanas, constituyen otro grupo extenso de lectinas que se ha caracterizado en detalle junto con el de las lectinas de origen vegetal. En 1974 se publicó por primera vez la descripción de una lectina de mamífero, el receptor de asialoglicoproteínas del hígado de conejo (Fersht, 1987) (**Figura 1.2B**). Sorprendentemente, tiempo después se demostró que la primera lectina animal había sido descubierta con

anterioridad, en 1853, cuando Charcot y Robin observaron los cristales ahora denominados de Charcot-Leyden en tejidos de pacientes enfermos de leucemia (Charcot y Robin, 1853). Estos cristales estaban realmente formados en una gran proporción por una proteína pura, la galectina-10 (Swaminathan *et al.*, 1999). Las lectinas de organismos invertebrados fueron caracterizadas en detalle hasta finales de la década de los ochenta, momento en el que los estudios de las lectinas procedentes de vertebrados tomó mayor relevancia. A diferencia de lo que ocurre con las lectinas de plantas, las lectinas de origen animal exhiben elevada similitud estructural, incluso entre especies alejadas evolutivamente. La primera clasificación de lectinas animales fue propuesta por Drickamer (Drickamer, 1988; Dodd y Drickamer, 2001), basándose en el tipo de plegamiento que presentaban los dominios de unión a carbohidratos o CRD (del inglés *Carbohydrate Recognition Domain*). Actualmente, se han establecido más de 12 de familias estructurales, siendo las más importantes las lectinas de tipo C<sup>23</sup> (dependientes de calcio), las galectinas (lectinas específicas de  $\beta$ -galactósidos) y las siglecs (proteínas de la superficie celular que reconocen ácido siálico).

### 1.2.3. Lectinas de microorganismos

Aunque con un comienzo más tardío, las lectinas provenientes de microorganismos (protozoos, bacterias y virus) han empezado también a ser objeto de importantes investigaciones. Además, gracias a que se está descubriendo más acerca de su participación en nuevos procesos, como la patogenicidad, que no son comunes a los descritos anteriormente para las lectinas de origen vegetal o animal, suscitan un gran interés.

En relación con el reino protista, se han caracterizado principalmente lectinas de organismos patógenos como *Plasmodium falciparum*, causante de la malaria, y *Trypanosoma cruzi*, responsable de la enfermedad de Chagas (Ward, 1997). Aunque la mayoría son citosólicas o de superficie, en *Trichomonas mobilensis* se han aislado lectinas específicas de ácido N-acetilneuramínico del medio extracelular (Babál *et al.*, 1994).

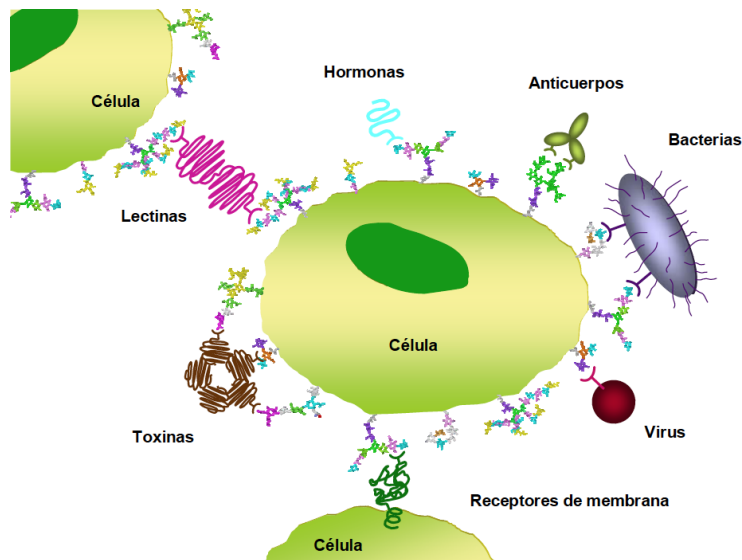
Con respecto a las lectinas de origen bacteriano, tanto las bacterias Gram negativas como las Gram positivas producen lectinas de amplia especificidad que suelen expresarse en su superficie, encontrándose asociadas a fimbrias y pelos (**Figura 1.2C**). Aunque no es frecuente, algunas lectinas bacterianas se expresan intracelularmente, como las descritas en *Pseudomonas aeruginosa* (Gilboa-Garber *et al.*, 1997).

Por último, las lectinas de origen viral se encuentran expuestas en la superficie de las partículas virales, pues están directamente implicadas en el reconocimiento de las células del huésped. Algunas de éstas, además, presentan actividad hemaglutinante, como la

hemaglutinina (HA) del virus de la gripe. Se han descrito lectinas en la práctica totalidad de tipos de virus, incluyéndose el virus del herpes simple (VHS), de la hepatitis C (VHC) o de la inmunodeficiencia humana (VIH) (Sharon y Lis, 2007).

### 1.3 ¿Qué hacen las lectinas?

Numerosos procesos biológicos transcurren mediante el reconocimiento de azúcares por lectinas en alguna de sus etapas. En ellos, las lectinas actúan *de facto* como vehículos en la transferencia de información en los sistemas biológicos, llevando a cabo sus tareas mediante la interacción con glicoproteínas, glicolípidos y/o mono y oligosacáridos. Generalmente, la función que van a desempeñar va a depender de su localización en la célula (**Figura 1.3**). Así, por ejemplo, las lectinas intracelulares van a estar involucradas en el tráfico de proteínas. Las lectinas asociadas a la superficie celular, sin embargo, jugarán un papel importante en procesos de adhesión, comunicación celular o reconocimiento de organismos patógenos. Las lectinas que se secretan al espacio extracelular podrán actuar como toxinas, como las descritas en algunas plantas, o estar asociadas a la capacidad invasiva de ciertos tumores, como en los glioblastomas (Camby *et al.*, 2006; Puchades *et al.*, 2007). Recientemente, se ha observado que algunas lectinas se localizan en el núcleo celular, pero su función aún no ha sido esclarecida del todo (Ahmed *et al.*, 2004; Lannoo *et al.*, 2006).



**Figura 1.3.** Diferentes interacciones proteína-azúcar que pueden tener lugar para mediar procesos celulares de distinta naturaleza y complejidad.

La participación de lectinas en diversos procesos biológicos las ha convertido en un importante objeto de estudio por parte de las industrias biomédica y biotecnológica. No

obstante, el desarrollo de aplicaciones basadas en la función de las lectinas va a estar siempre ligado al conocimiento detallado de las características tanto estructurales como termodinámicas del reconocimiento de los carbohidratos, que completan la comprensión global del sistema biológico.

### 1.3.1. Lectinas de acción intracelular

Las proteínas de reserva que se encuentran en las plantas, sobre todo en sus semillas, son una fuente rica de lectinas intracelulares. Las lectinas aparecen en los cuerpos proteicos junto con otras proteínas a las que se unen mediante interacciones de tipo lectina-azúcar y proteína-proteína. Aunque representan un elevado porcentaje de las lectinas en plantas, la función que desempeñan no se ha dilucidado aún. Se cree que algunas de estas lectinas podrían actuar como un puente entre los cuerpos proteicos y la membrana, organizando su contenido para su posterior disponibilidad por parte de la célula durante la maduración de las semillas (Wenzel *et al.*, 1993).

Por otro lado, los mecanismos por los cuales las lectinas regulan la actividad de algunas enzimas hidrolíticas, como las fosfatasas del germen de trigo y la de patata, han sido caracterizados en detalle (Rudiger y Gabius, 2001). En algunos casos, se han encontrado funciones adicionales además de la mera unión a carbohidratos. Por ejemplo, la cadena B de la ricina muestra actividad lipolítica, que aparentemente contribuye también a su toxicidad (Helmy *et al.*, 1999). Además de estas funciones, existen otras que comparten las lectinas de plantas y animales, relacionadas sobre todo con la regulación de diferentes procesos citoplasmáticos.

Una de las primeras lectinas que se identificó como clave en el tráfico intracelular fue el receptor de manosa-6-fosfato, cuya función es dirigir las enzimas lisosomales desde el sitio de síntesis en el retículo endoplasmático rugoso hacia los lisosomas, donde van a llevar a cabo su función (Dahms *et al.*, 1989). A finales de los años ochenta, la evidencia acumulada demostró que algunas lectinas también pueden actuar como chaperonas. Un ejemplo clave de esta función es el sistema de control de calidad de la síntesis de glicoproteínas N-glicosiladas. La calnexina, una lectina integral de membrana del retículo endoplasmático, funciona en paralelo con la calreticulina, su homólogo soluble, tratando de asegurar el correcto plegamiento de las glicoproteínas que, en última instancia, se dirigirán a la superficie celular (Trombetta y Helenius, 1998).

Hoy en día es sabido que el reconocimiento de azúcares por lectinas intracelulares juega un papel clave en numerosos procesos biológicos, como la regulación de la síntesis de

proteínas, la diferenciación celular o la activación de cascadas de señalización que conducen, por ejemplo, a la apoptosis (muerte celular programada) (Lichtenstein y Rabinovich, 2013).

### 1.3.2. Lectinas de la superficie celular

El reconocimiento de carbohidratos por lectinas subyace en los mecanismos de fijación observados en numerosas bacterias, virus, hongos y protozoos (Vasta, 2009). En la superficie de muchas bacterias, por ejemplo, hay lectinas (adhesinas) que facilitan la adhesión y posterior colonización al unirse a glicoconjugados de la membrana de la célula huésped. En algunos casos, la especificidad de unión explica el tropismo hacia un tejido concreto del organismo (Petukhov *et al.*, 1999). Por ejemplo, la infección del tracto urinario por un serotipo específico de *Escherichia coli* depende de la unión a manosa.

En la década de los años cincuenta, se demostró por primera vez que el virus de la gripe reconoce azúcares específicos en la célula huésped. La hemaglutinina del virus se une específicamente a los residuos de ácido siálico de los oligosacáridos de la superficie celular (Weis *et al.*, 1988). Por el contrario, en otros casos, son las lectinas de la célula huésped las que están involucradas en la unión del virus. Por ejemplo, las lectinas DC-SIGN y L-SIGN, localizadas en las células dendríticas y las células endoteliales del hígado y nódulos linfáticos, entre otros tejidos, respectivamente, pueden mediar la adhesión de los virus VIH-1, citomegalovirus humano, VHC y virus Ébola. El reconocimiento tiene lugar a través de la unión de estructuras ricas en manosa presentes en la superficie de los virus (Álvarez *et al.*, 2002).

Las lectinas de superficie son también importantes para otros procesos fisiológicos de relevancia como, por ejemplo, la fecundación y el desarrollo embrionario, así como en la formación de tumores y la diseminación de células cancerígenas desde el tumor inicial a otros tejidos. También juegan un papel crucial en la respuesta inmunológica, participando en los procesos de inflamación y defensa del organismo. De hecho, juegan un papel crítico en la definición de la identidad del organismo porque permiten distinguir entre los azúcares propios y los ajenos, asegurando así una respuesta inmunológica eficaz. Algunas lectinas endógenas se unen a carbohidratos presentes en la superficie de microorganismos patógenos, provocando su encapsulación y muerte, proceso llamado lectinofagocitosis (Ofek y Sharon, 1988). Entre dichas lectinas cabe citar el receptor de macrófagos específico de manosa, la dectina-1, específica de  $\beta(1\rightarrow3)/(1\rightarrow6)$  glucanos presentes en hongos. La maduración tímica, la activación, migración y apoptosis de células T está también regulada por la

interacción específica de lectinas, como las selectinas y las galectinas. La expresión de estas lectinas se regula de forma positiva o negativa en función de la fase de la respuesta inmunológica (van Kooyk y Rabinovich, 2008). Las selectinas también regulan la unión de leucocitos a las células endoteliales, iniciando la fase de diapédesis que lleva finalmente a su extravasación.

### 1.3.3. Lectinas de acción extracelular

Gran parte de las lectinas bacterianas que son secretadas al medio extracelular actúan como potentes toxinas, siendo los principales agentes moleculares responsables de la patogenicidad del organismo. Como ejemplos se pueden citar a la toxina A de *Clostridium difficile*, la toxina del cólera (*Vibrio cholerae*) o las enterotoxinas de *E. coli*, que actúan sobre las células epiteliales del intestino.

En el medio natural, la función de algunas lectinas extracelulares como potentes toxinas frente a predadores es clave para la supervivencia. Esta función fue inicialmente identificada en lectinas como la ricina o la aglutinina de la judía verde (*Phaseolus vulgaris*) (PHA) (Liener, 1986). Esto hace que el consumo de ciertas plantas deba hacerse con precaución, ya que, por ejemplo, si la judía verde no está bien cocinada puede producir irritación gastrointestinal. Algunas lectinas también actúan como un sistema de defensa frente a hongos, interfiriendo en la síntesis de la pared celular, alterando, por tanto, su crecimiento (Sharon y Goldstein, 1998). Mecanismos de defensa similares se han descrito en animales, siendo las colectinas y las ficolinas los agentes principales contra la infección (Lu *et al.*, 2002).

No todas las lectinas de acción extracelular forman parte de mecanismos de defensa. En procesos simbióticos entre leguminosas y bacterias fijadoras de nitrógeno del género *Rhizobium*, una lectina presente en las raíces de la planta reconoce específicamente los lipooligosacáridos, también denominados factores de nodulación (Nod), facilitando el anclaje de las bacterias (Kalsi y Etzler, 2000).

A semejanza de las proteínas fúngicas y bacterianas que se encargan de la degradación de la pared celular, hay descrito un tipo de lectinas de plantas de acción extracelular que aparecen asociadas a enzimas hidrolíticas. El módulo lectina actúa como guía de la enzima, dirigiéndola hacia el sustrato correspondiente. Ejemplos de estas proteínas son la  $\beta$ -galactosidasa de fresa y la celulasa de tomate (Boraston *et al.*, 2004).

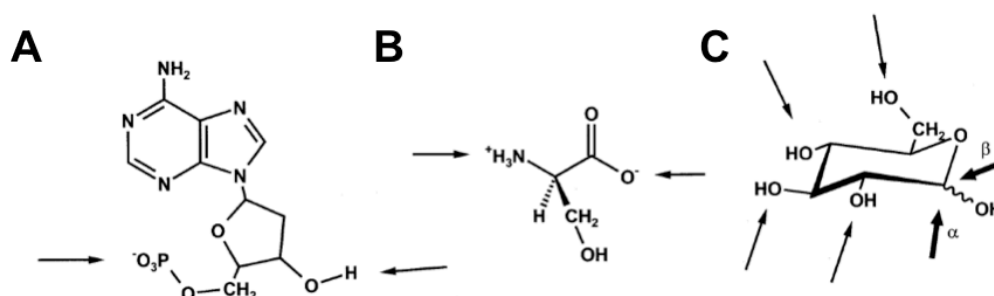
## 1.4 El glicocódigo

Como se ha visto en los apartados anteriores, las lectinas están involucradas en multitud de procesos biológicos. La razón principal es que en estos procesos está implicada una gran diversidad de carbohidratos o azúcares, los cuales actúan como ligandos de dichas lectinas. Los carbohidratos son moléculas cuya importancia en un contexto biológico puede equipararse a la de los ácidos nucleicos o las proteínas en cuanto que soporte a la información biológica (Roseman, 2001). No obstante, entre ellos hay notables diferencias, tales como la complejidad en la relación entre la estructura de los carbohidratos y el genoma, en cuanto que, contrariamente a ácidos nucleicos y proteínas, los carbohidratos no son codificados por éste (Gupta y Surolia, 2012). Los carbohidratos están presentes mayoritariamente formando glicoconjugados<sup>17</sup>, es decir, se encuentran incorporados a proteínas o lípidos mediante reacciones de modificación postraducciona, llevadas a cabo por glicosiltransferasas muy específicas cuya actividad depende de varios factores ambientales, entre los que se encuentran la concentración de azúcar, la localización celular, el metabolismo, los niveles de expresión de los *glicogenes*, etc. Por ejemplo, prácticamente todas las proteínas nucleares y de unión a ADN, enzimas citoplasmáticas involucradas en la regulación metabólica y algunas proteínas mitocondriales poseen residuos de Ser/Thr modificados con *O*- $\beta$ -N-acetilglucosamina (*O*-GlcNAc) (Cummings y Pierce, 2014). Esto conlleva, por tanto, a que, a diferencia del genoma, el glicoma<sup>18</sup> sea variable y dependa a su vez fuertemente de distintos factores no genéticos.

Además del aspecto biológico, la estructura de los carbohidratos puede ser notablemente compleja debido a que las unidades básicas que los componen ofrecen una extraordinaria versatilidad para combinarse. Cuando se compara la capacidad “comunicativa” de los carbohidratos con la de los nucleótidos o aminoácidos, es decir, el número de “palabras” (oligosacáridos) que pueden construirse con las “letras” (monosacáridos) disponibles, ésta es varios órdenes de magnitud superior: por ejemplo, el número teórico de hexámeros posibles sería de 4.096 hexanucleótidos,  $6,4 \times 10^7$  hexapéptidos y  $1,44 \times 10^{15}$  hexasacáridos (Laine, 1997). Esta complejidad es debida a la contribución de diferentes factores: (i) posibilidad de dos configuraciones anoméricas<sup>4</sup> ( $\alpha$  o  $\beta$ ), (ii) formación de enlaces glicosídicos a partir de diferentes posiciones del azúcar (1 $\rightarrow$ 1, 2, 3, 4, 5, 6 para una hexopiranososa), (iii) cambios en el tamaño del anillo (piranososa<sup>30</sup> o furanososa<sup>15</sup>), (iv) modificaciones covalentes como acetilaciones, fosforilaciones o sulfataciones, y (v) formación de polisacáridos ramificados (**Figura 1.4**) (Gabiús *et al.*, 2011). Por otra parte, mientras la conformación de silla



de un monosacárido es relativamente rígida, el ángulo de torsión alrededor del enlace glicosídico puede variar, de modo que una estructura primaria bien definida puede adoptar diversas conformaciones en solución en las que la orientación relativa de los monosacáridos es diferente.



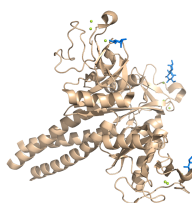
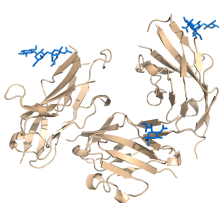
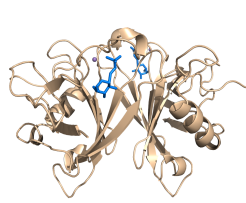
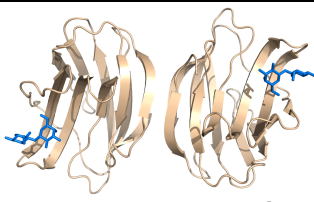
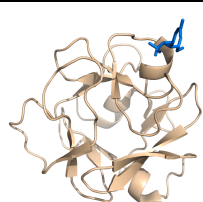
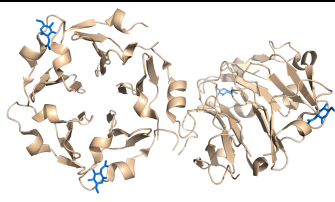
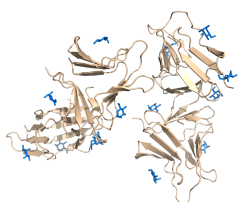
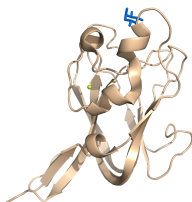
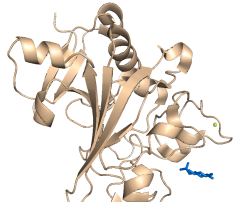
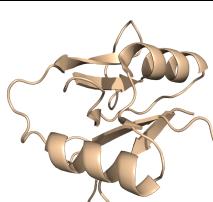
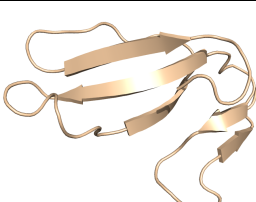
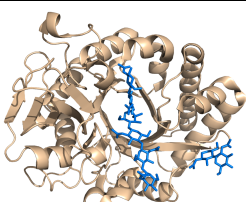
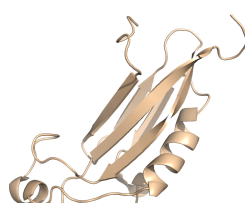
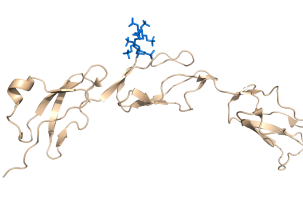
**Figura 1.4.** Puntos de anclaje para la formación de oligómeros en los diferentes tipos de biomoléculas. Los nucleótidos (A) y los aminoácidos (B) solo pueden formar oligómeros lineales, mientras que los monosacáridos (C) no tienen ningún tipo de restricción *a priori* para formar estructuras ramificadas.

Cada carbohidrato adquiere un significado en función de las unidades (monosacáridos) que lo integran. Un aspecto característico es la orientación de los grupos hidroxilo, los más numerosos, ya que su disposición espacial va a ser crucial para el establecimiento de enlaces de hidrógeno direccionales con los aminoácidos presentes en el sitio de unión de la lectina. Además, la orientación de los anillos también va a ser un factor clave, puesto que al ser más hidrofóbicos podrán interactuar mediante fuerzas de van der Waals con aminoácidos apolares y mediante interacciones C-H/ $\pi$ <sup>20</sup> con residuos aromáticos. Teniendo en cuenta que muchos azúcares sólo se diferencian en la orientación de uno o dos grupos hidroxilo en torno a carbonos quirales, se puede decir que la epimerización<sup>11</sup> es un factor importante en el biorreconocimiento (Drickamer, 1997).

## 2. ¿Qué nos cuentan las estructuras de las lectinas?

Para reconocer a los diferentes azúcares presentes en un organismo, las lectinas adoptan un plegamiento específico que dispone a los aminoácidos responsables de la unión en una cavidad o región concreta de la macromolécula que permite la unión del ligando correspondiente. La rica variabilidad estructural de los azúcares reconocidos por las lectinas se traduce a su vez en la existencia de un gran número de estructuras en éstas, capaces de acomodar gran variedad de carbohidratos, desde monosacáridos sencillos hasta polisacáridos de elevada complejidad. Actualmente, se ha establecido una clasificación que distinga entre, al menos, 14 plegamientos distintos dentro de la superfamilia de las lectinas

(Gabiús *et al.*, 2011) (**Figura 1.5**), cada uno con su propia diversidad filogenética (Houzelstein *et al.*, 2004).

Tipo C	Tipo I	Tipo P
 1PWB	 1QFO	 1M6P
Sándwich $\beta$	Trébol $\beta$	Hélice enrollada
 1GZW	 1DQO	 3KIH
Prisma $\beta$ II	Barril $\beta$ tipo remolino	Tipo fibrinógeno
 1KJ1	 1K12	 1JC9
Módulo eslabón	Tipo heveína	Barril $(\alpha/\beta)_8$
 2PF5	 1UG4	 1HJW
Plegamiento $\alpha/\beta$	Dominio SCR	<p><b>Figura 1.5.</b> Algunos ejemplos que muestran la diversidad de plegamientos dentro de la superfamilia de las lectinas. En cada caso se indica el código PDB de la estructura representada.</p>
 2JX9	 2V8E	

Desde que en los años setenta comenzase el estudio estructural de las lectinas, el número de estructuras caracterizadas ha ido creciendo rápidamente, tanto de lectinas en

su forma libre como formando algún tipo de complejo. La mayoría de las estructuras se han obtenido mediante cristalografía de rayos X<sup>8</sup>, llegándose a un nivel de resolución comprendido entre 1,80-2,20 Å ([www.glyco3d.cermav.cnrs.fr](http://www.glyco3d.cermav.cnrs.fr)). Hasta la fecha, las lectinas de tipo C constituyen el grupo de lectinas más estudiado, conjuntamente con el de las galectinas (Drickamer, 1999). La descripción de las estructuras tridimensionales no sólo sirve para entender mejor el mecanismo de la interacción lectina-azúcar, sino que permite establecer relaciones evolutivas entre las lectinas presentes en diferentes tipos de seres vivos. Es más, la comparación de los determinantes de la unión de los carbohidratos identificados en las estructuras de los complejos lectina-azúcar permite estudiar cómo evolucionan los procesos de reconocimiento a medida que aumenta la complejidad tanto estructural como química de los carbohidratos presentes en los diferentes organismos (Drickamer, 1997).

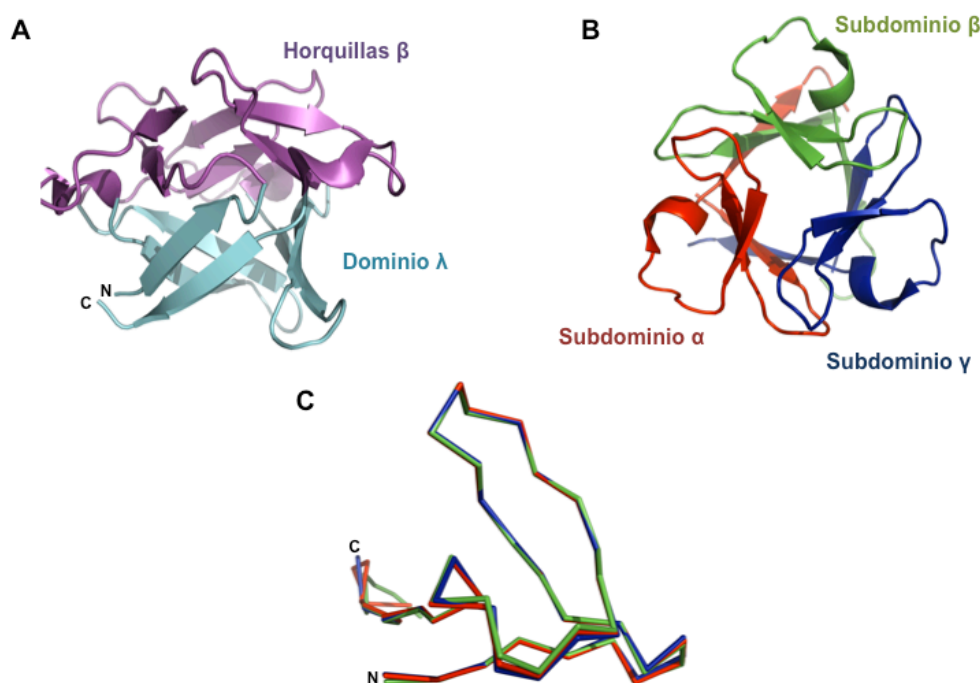
En el presente trabajo, las proteínas objeto de estudio se engloban dentro del grupo de lectinas con estructura tipo trébol  $\beta$ <sup>37</sup>. Esta estructura es bastante común en diferentes grupos de lectinas y fue identificada por primera vez en el inhibidor de tripsina de la soja (Sweet *et al.*, 1974).

## 2.1. La estructura tipo trébol $\beta$

La estructura trébol  $\beta$  está constituida por tres subdominios de unos 50 residuos cada uno, frecuentemente denominados  $\alpha$ ,  $\beta$  y  $\gamma$ , desde el extremo N al C-terminal, cada uno de ellos formado por cuatro hebras  $\beta$  antiparalelas. Los tres subdominios son estructuralmente parecidos y muestran cierta similitud de secuencia entre ellos. El plegamiento global da lugar a una estructura globular, que presenta una pseudosimetría<sup>32</sup> interna de orden tres (Loris, 2002), formada por un pequeño barril  $\beta$  de seis hebras (también conocido como dominio  $\lambda$ ) con una tapadera formada por un triplete de horquillas  $\beta$  en uno de sus lados (**Figura 1.6**). En principio, cada uno de los subdominios tiene la capacidad de unir azúcares, por lo que un trébol  $\beta$  podría tener hasta tres sitios definidos de unión de azúcares ( $\alpha$ ,  $\beta$  y  $\gamma$ ).

El plegamiento de tipo trébol  $\beta$  fue identificado por primera vez como un dominio de reconocimiento de carbohidratos en la proteína ricina (Montfort *et al.*, 1987; Rutenber *et al.*, 1991) (**Figura 1.7**). La ricina es una glicoproteína heterodimérica (AB) aislada de las semilla del ricino (*R. communis*), cuya actividad citotóxica ha sido estudiada en detalle. Pertenecce a la familia de las proteínas inactivadoras de ribosomas o RIP (del inglés *Ribosome Inactivating Proteins*) y su interés se ha centrado, sobre todo, en su posible aplicación tera-

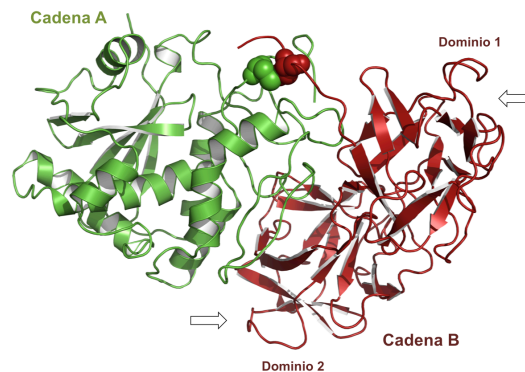
péutica. La cadena B de la ricina tiene la capacidad de unir azúcares y está compuesta por dos dominios globulares, siendo cada uno de ellos un trébol  $\beta$ . Aunque la ricina tendría potencialmente seis sitios de unión de azúcares, sólo presenta dos sitios operativos, localizados en el subdominio  $1\alpha$  de la zona N-terminal y en el  $2\gamma$  de la C-terminal. Hoy día se sabe que esto se debe a que son los únicos sitios que presentan residuos aromáticos (Trp-37 y Tyr-248, respectivamente) capaces de establecer interacciones de apilamiento con el anillo de piranosa del azúcar.



**Figura 1.6.** Representación tridimensional de la estructura trébol  $\beta$  de la proteína con código PDB 304D. En **A** se muestra el dominio  $\lambda$  o barril  $\beta$  (cian) y el triplete de horquillas  $\beta$  (magenta). En **B**, con una vista cenital del trébol  $\beta$ , se distinguen los tres subdominios:  $\alpha$  (rojo),  $\beta$  (verde) y  $\gamma$  (azul), cuya superposición (**C**) demuestra que son prácticamente idénticos.

Después de la resolución estructural de la ricina, el plegamiento de tipo trébol  $\beta$  se ha identificado en otras proteínas de plantas, como la lectina amarantina del amaranto (*Amaranthus caudatus*) (Transue *et al.*, 1997). Por otro lado, en animales se han encontrado dos familias de proteínas que adoptan este mismo tipo de plegamiento: los factores de crecimiento de fibroblastos (FGF) (DiGabriele *et al.*, 1998) y las lectinas tipo R, entre las que se encuentran, por ejemplo, el dominio receptor de manosas rico en cisteínas (cys-MR) (Liu *et al.*, 2000). En la última década y gracias a iniciativas conjuntas de los consorcios de genómica estructural, se han caracterizado más lectinas con esta misma estructura en bacterias (Tanaka *et al.*, 2009) y hongos. Actualmente y de acuerdo a la clasificación Pfam

(Punta *et al.*, 2012), las proteínas con estructura tipo trébol  $\beta$  comprenden 15 familias distintas.



**Figura 1.7.** Estructura cristalográfica del heterodímero AB de la ricina (PDB 2AAI). En *verde* se muestra la cadena A con actividad RIP, y en *rojo* la cadena B o módulo lectina. Las cisteínas que forman el puente disulfuro entre las dos cadenas están representadas con el modelo de *bolas*. Las flechas señalan las posiciones de los sitios de unión a carbohidratos presentes en los dos dominios de la cadena B.

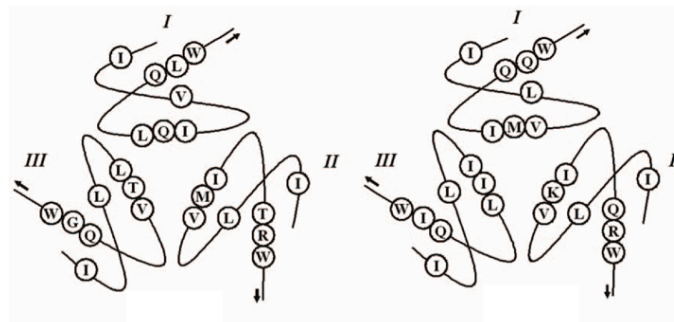
## 2.2. Evolución del plegamiento trébol $\beta$

Un gran número de proteínas en la naturaleza presenta elementos de simetría interna en su estructura, como ocurre con las lectinas con estructura tipo trébol  $\beta$ . La hipótesis más plausible que trata de explicar cómo se originan estas proteínas es la de la duplicación y fusión génica. Concretamente, esta hipótesis contempla la ocurrencia de un fenómeno de duplicación de un gen ancestral, que codificaría un pequeño péptido ( $\sim 40$ -60 aminoácidos), y otro de fusión, que originaría un gen responsable de una proteína de tipo trébol  $\beta$ , dotado de simetría interna (Tateno *et al.*, 1997). A partir de esta hipótesis se han considerado dos escenarios en función de cuál es la unidad evolutiva en los tréboles  $\beta$ : (1) el modelo de *arquitectura emergente*, en el que los subdominios constituyen la unidad evolutiva y el trébol aparece como propiedad emergente resultado de la duplicación/fusión; y (2) el modelo de la *arquitectura conservada*, en el que el trébol  $\beta$  propiamente dicho es la unidad evolutiva, apareciendo incluso antes de los acontecimientos de duplicación/fusión, mediante procesos de oligomerización de los péptidos ancestrales.

En el caso particular del plegamiento tipo trébol  $\beta$ , se ha observado recientemente que éste sigue un modelo evolutivo de arquitectura conservada (Lee y Blaber, 2011). Tomando como punto de partida la secuencia que codifica un solo subdominio optimizado en términos de simetría, el péptido resultante fue capaz de asociarse y formar un homotrímero, esto es, un trébol  $\beta$  completo. A su vez, el producto de un primer acontecimiento de fusión (péptido con dos subdominios) también oligomerizó formando dos tréboles  $\beta$  comple-

tos. Este resultado indica la validez del modelo de estructura conservada sobre el anteriormente propuesto de estructura emergente por Mukhopadhyay (Mukhopadhyay, 2000). Además, ofrece una vía para comprender la aparición de un nuevo nivel de complejidad estructural, el de las proteínas con simetrías internas, desde genomas aparentemente más simples.

Otro de los aspectos de la aparición del plegamiento trébol  $\beta$  gira en torno a la conservación de la estructura primaria. En este sentido, aunque los análisis de secuencias sugieren que todos los tréboles  $\beta$  evolucionaron desde un ancestro común homotrimérico (Murzin *et al.*, 1992; Mukhopadhyay, 2000; Ponting y Russell, 2000), estudios más recientes y a gran escala encuentran resultados que apuntan a que no existe tal ancestro único, sino más bien sugieren que los procesos de repetición y fusión de subdominios es un proceso activo y recurrente desde un punto de vista evolutivo, habiendo ocurrido numerosas veces (Broom *et al.*, 2012).



**Figura 1.8.** Diagrama donde se muestran los cuatro motivos FTR conservados en los dos dominios con estructura trébol  $\beta$  de la cadena B de la ricina. Las flechas indican la dirección de las hebras  $\beta$ .

Aunque la estructura primaria no se conserva, en la mayoría de los tréboles  $\beta$  existen ciertos motivos relativamente frecuentes. Así, Hazes detectó en la secuencia de la cadena B de la ricina que el motivo QXW (donde la X es cualquier aminoácido) se encontraba repetido en cada uno de los subdominios (Hazes, 1996). Más tarde, se llegó a la conclusión de que este motivo es una característica común presente en prácticamente todos los tréboles  $\beta$  (Murzin *et al.*, 1992). Mediante un análisis exhaustivo de la secuencia de aminoácidos de diferentes citotoxinas de plantas vegetales con estructura tipo trébol  $\beta$ , se llegó a la conclusión de que existen cuatro motivos comunes conservados en todos los subdominios:  $(I)_3$ ,  $(L/M/V)_3$ ,  $([I/L/V]X[I/L/M])_3$  y  $(QXW)_3$ ; donde la X siempre indica cualquier residuo. Los residuos de los cuatro motivos repetidos por partida triple (FTR, por sus siglas en inglés, *Four Three-Fold Repetitive Motifs*) se distribuyen simétricamente en la estructura tridimen-

sional de las proteínas, localizándose en posiciones clave (**Figura 1.8**) y siempre en zonas enterradas con menos de un 25% de accesibilidad al solvente. El primer motivo se localiza en lo alto del barril  $\beta$ , el cuarto en la parte de en medio y los dos que quedan en la base (Feng *et al.*, 2010).

### 3. Interacciones lectina-azúcar

Gracias a técnicas estructurales de alta resolución como la cristalografía de rayos X y a técnicas termodinámicas como la calorimetría de titulación isotérmica, se ha avanzado en el estudio detallado de los sitios de unión y en la comprensión de los mecanismos de interacción con los azúcares. Aunque las lectinas exhiben gran variedad de estructuras, todas comparten ciertas características relativas a los sitios de unión: en general, estos aparecen en hendiduras poco profundas en la superficie de la proteína y muestran una relativa baja afinidad por azúcares sencillos (mono y disacáridos). Como se describió anteriormente, la selectividad de las lectinas se establece mediante la formación de enlaces de hidrógeno con los grupos hidroxilo del azúcar e interacciones de van der Waals y apilamientos que involucran aminoácidos aromáticos. Los incrementos de afinidad se observan con azúcares más complejos que acomodan parte de sus residuos en el entorno del sitio de unión, que se vería por lo tanto extendido. Por último, es importante no olvidar que las moléculas de agua juegan un papel clave en la fijación de los ligandos (Weis y Drickamer, 1996).

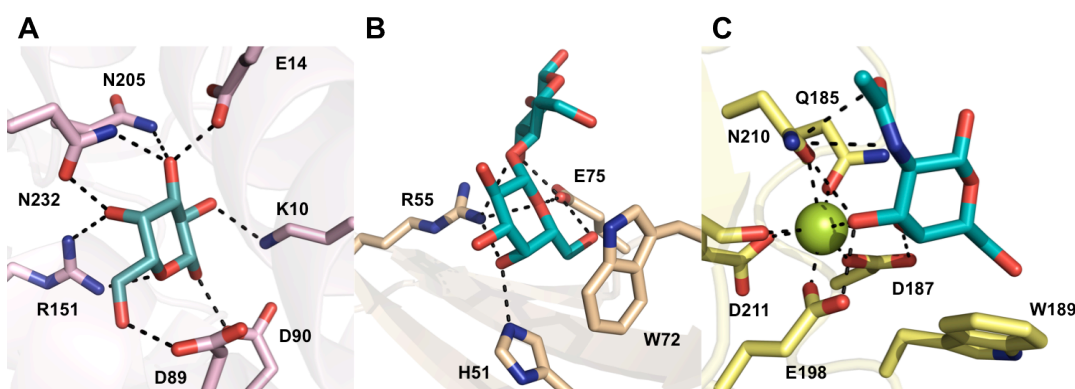
A diferencia de otras proteínas, las lectinas no experimentan cambios conformacionales significativos por la unión del ligando. Se han observado pequeños desplazamientos de residuos localizados en el sitio de unión o próximos a éste. Estas pequeñas reorganizaciones de estructura en el entorno del sitio de unión se pueden identificar y evaluar mediante cambios en la fluorescencia intrínseca de la proteína o por resonancia magnética nuclear (RMN). Sin embargo, el estudio de estos cambios tan sutiles mediante cristalografía de rayos X es más complicado, puesto que el empaquetamiento de las moléculas en los cristales puede restringir la movilidad de los residuos.

#### 3.1. Principales interacciones lectina-azúcar

El principal tipo de interacción que se establece entre los azúcares y las lectinas es el enlace de hidrógeno. Los grupos hidroxilo tienen la capacidad de actuar simultáneamente como dadores y aceptores en este tipo de interacciones (**Figura 1.9A**), contribuyendo significativamente a la especificidad. El oxígeno, que presenta hibridación  $sp^3$ , posee una geometría



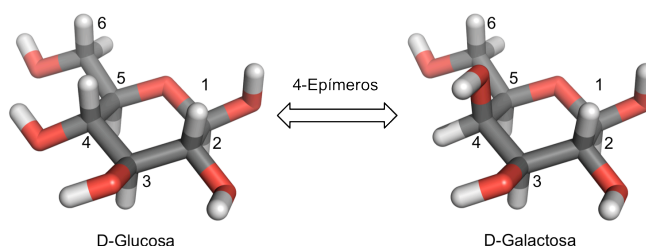
tetraédrica que le permite interaccionar con diferentes residuos a la vez. Estos residuos suelen ser aminoácidos polares que presentan una geometría plana en su cadena lateral (Asn, Asp, Gln, Glu, His y Arg) (Petukhov *et al.*, 1999). En la mayoría de los casos, en el sitio de unión se observa un residuo ácido (Asp o Glu), que actúa como aceptor de uno o dos grupos hidroxilo del azúcar, así como aminoácidos con grupos amida (Asn o Gln) que actúan como dadores. Por otro lado, la infrecuente participación de residuos con grupos hidroxilo (Ser, Thr) en los sitios de unión se ha relacionado con el hecho de que su libertad rotacional no permitiría proporcionar especificidad a la interacción con el azúcar. En ocasiones, también participa el oxígeno del anillo del azúcar (igualmente con hibridación  $sp^3$ ) en interacciones con la proteína. Este oxígeno tiene dos pares de electrones libres que le permiten funcionar como aceptor de hidrógeno. Puesto que este átomo de oxígeno es común a la práctica totalidad de azúcares, no es un determinante de la especificidad.



**Figura 1.9.** Ejemplos de los tres tipos de interacciones más frecuentes entre lectinas y azúcares. **A.** Red de enlaces de hidrógeno establecida entre la proteína de unión a L-arabinosa de *E. coli* (5ABP) y un residuo de D-galactosa. **B.** Apilamiento entre el anillo de galactosa de una molécula de lactosa y el Trp-72 del sitio de unión de azúcares de la galectina fúngica CGL2 (1UCL). **C.** Sitio de unión de la lectina sérica de unión a manosa de rata (1BCH) en donde aparece un átomo de calcio en contacto directo con la lectina y la molécula de N-acetil-D-galactosamina.

La especificidad de las lectinas por una clase de azúcares viene generalmente determinada por los enlaces de hidrógeno que se establecen con los grupos hidroxilo que diferencian a unos azúcares de otros. Por ejemplo, los grupos hidroxilo en posición C4 de la glucosa y la galactosa presentan diferente orientación (**Figura 1.10**), por lo que se consideran clave para la especificidad de las lectinas. Una vez que estos enlaces se han formado, el resto de interacciones que tienen lugar tanto por enlace de hidrógeno como por otro tipo de interacciones intermoleculares contribuyen únicamente a la estabilidad, aumentando, de esta manera, la intensidad de la unión (Drickamer, 1997).





**Figura 1.10.** Representación esquemática donde se muestra la existencia de estereoisomería o isomería de posición del grupo OH en torno al átomo de carbono C4 entre los monosacáridos D-glucosa y D-galactosa. Esta diferencia en la disposición espacial del grupo hidroxilo hace que ambas moléculas (4-epímeros) sean reconocidas de distinta forma por las lectinas.

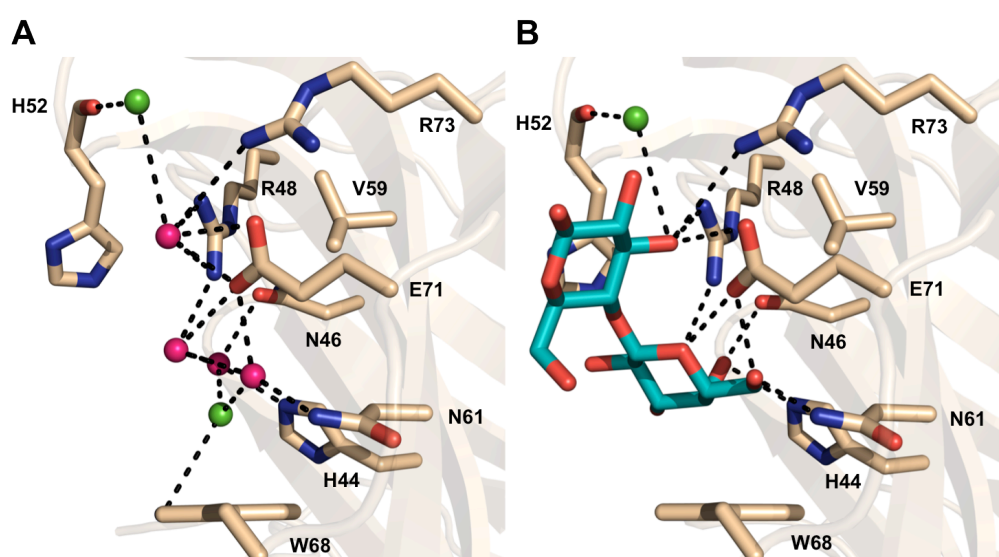
Junto con los enlaces de hidrógeno, las interacciones apolares son las que más contribuyen a la interacción lectina-azúcar. Los carbohidratos suelen presentar zonas hidrofóbicas formadas por hidrógenos alifáticos y átomos de carbono. Estas regiones interactúan frecuentemente con uno o más anillos aromáticos de la lectina (Trp, Tyr o, menos frecuentemente, Phe) mediante interacciones C-H/  $\pi$  (Levitt y Perutz, 1988) (**Figura 1.9B**). Además de la complementariedad geométrica entre el anillo de la molécula de azúcar y los residuos aromáticos, la carga negativa de la nube electrónica  $\pi$  del anillo aromático intensifica la carga parcial de los hidrógenos alifáticos del ligando, lo que da lugar a una interacción más fuerte. Dos requisitos esenciales para que se den estas interacciones son que el azúcar y los anillos aromáticos estén dispuestos cara a cara y que no disten más de 4 Å entre sí. Al no quedar expuestas al solvente las zonas hidrofóbicas del azúcar ni el anillo aromático, la contribución energética a la unión de este tipo de interacciones es bastante elevada (Toone, 1994; Weis y Drickamer, 1996).

Las modificaciones covalentes en las moléculas de azúcar pueden aumentar la afinidad de la unión y actuar como determinantes de la especificidad. Por ejemplo, el grupo acetamido de la N-acetilactosamina, que presenta geometría plana, tiene la posibilidad de formar nuevos enlaces de hidrógeno a través del grupo amida. Por otro lado, la presencia de grupos cargados (carboxilatos, sulfatos o fosfatos) puede facilitar el establecimiento de fuertes interacciones electrostáticas.

Por último, en algunos casos se han encontrado cationes divalentes como  $\text{Ca}^{2+}$  y  $\text{Mn}^{2+}$  implicados en el reconocimiento, ya sea indirectamente, configurando el sitio de unión (por ejemplo, en lectinas de plantas leguminosas como la ConA), o por unión directa al carbohidrato, como ocurre en las lectinas de tipo C (**Figura 1.9C**). En ambos casos, la presencia de estos iones es necesaria para mantener la capacidad de unión a azúcares.

### 3.2. Papel del agua en las interacciones lectina-azúcar

Tal y como se observa en las estructuras de alta resolución, las moléculas de agua participan en la interacción lectina-azúcar, actuando de puente entre la proteína y el ligando (Tschampel y Woods, 2003). Estos enlaces de hidrógeno son también direccionales y pueden ser tan fuertes como los que se establecen directamente entre el azúcar y la lectina (Watson *et al.*, 1994). La comparación entre los CRD de diferentes lectinas demuestra que existen moléculas de agua en posiciones equivalentes, lo que evidencia que se trata de un elemento importante en el reconocimiento del azúcar.



**Figura 1.11.** Representación del sitio de reconocimiento de carbohidratos en el mutante C2S de la galectina-1 humana. **A.** Sitio de unión de la proteína libre, representando en *magenta* las aguas que ocupan las mismas posiciones que los oxígenos del azúcar y en *verde* las aguas que estabilizan a las anteriores. **B.** Sitio de unión del complejo con lactosa. La lactosa se posiciona de tal manera que sustituye a las aguas que preformaban el sitio de unión.

Cuando se compara la estructura de una lectina en su estado libre con la del complejo con un azúcar, se observa que en la primera hay moléculas de agua que mimetizan las interacciones que se establecen con el azúcar y estabilizan la orientación de los residuos proteicos involucrados en la unión del ligando (**Figura 1.11**). La unión del azúcar a la lectina transcurre con el desplazamiento de las moléculas de agua presentes en el sitio de unión (factor entrópico favorable), permitiendo así el establecimiento de las interacciones directas entre la lectina y el azúcar. Pero, por otro lado, se observan otras moléculas de agua que no son desplazadas y participan en la unión, por lo que se las puede considerar como una extensión de la propia estructura proteica, es decir, parte de la arquitectura del sitio de unión (Angulo *et al.*, 2011). Este fenómeno justifica en parte la no existencia de grandes

cambios conformacionales cuando el azúcar se une a la proteína. En general, las moléculas de agua contribuyen a aumentar la afinidad, pero no determinan la especificidad de la unión de los azúcares (Raschke, 2006).

Es interesante mencionar que si se calcula la distribución espacial de la probabilidad de albergar moléculas de agua en los sitios de unión (WS, por sus siglas en inglés, *Water Sites*), las posiciones que presentan mayor probabilidad son generalmente aquellas reservadas para los grupos hidroxilo del carbohidrato en el complejo (Di Lella *et al.*, 2007). De este modo, estudiando las propiedades de los WS en la proteína libre es posible predecir la configuración del ligando unido (Roldos *et al.*, 2011).

### 3.3. Extensión de los sitios de unión

Aunque las lectinas pueden presentar especificidad por un monosacárido en concreto, suelen ser más versátiles a la hora de acomodar otros residuos que acompañan al determinante inicial que media el reconocimiento en primer lugar. Este comportamiento se traduce, generalmente, en un incremento sustancial de la afinidad por el oligosacárido frente a los monosacáridos, lo cual es, por otro lado, esperable desde una perspectiva puramente termodinámica. De esta manera, surge el concepto de *sitio extendido* o *subsitio*<sup>35</sup> para hacer referencia a esas zonas de la superficie de la lectina que también interaccionan con el azúcar con la finalidad de incrementar la afinidad global por un oligosacárido. Por ejemplo, la lectina del guisante muestra alta afinidad por oligosacáridos que contienen residuos de fucosa unidos mediante enlace  $\alpha(1\rightarrow6)$  (Kornfeld *et al.*, 1981). La acomodación de residuos en los subsitios secundarios no solo está favorecida entálpicamente, sino que además la penalización entrópica es menor por el reducido número de grados de libertad conformacional del oligosacárido frente a las unidades de monosacárido individuales.

### 3.4. Multivalencia y cooperatividad

Muchas lectinas tienen más de un sitio de unión a azúcar en su estructura con características fisicoquímicas distintas. Normalmente, los sitios secundarios surgen por fenómenos de duplicación y, por ello, suelen ser similares al primero en lo que a la especificidad respecta (Weis y Drickamer, 1996). La presencia de múltiples sitios de unión independientes da lugar a incrementos en la afinidad cuando la lectina reconoce ligandos multivalentes<sup>24</sup>, como pueden ser oligosacáridos ramificados. El aumento de afinidad por la presencia de múltiples sitios de unión se denomina multivalencia<sup>26</sup> (Brewer *et al.*, 2002; Dam *et al.*, 2002).

Ésta aparece también en algunas lectinas con un único sitio de unión por subunidad pero que muestran un comportamiento asociativo y forman oligómeros en solución. Estos oligómeros muestran una afinidad mayor por el azúcar que las subunidades individuales (Earl *et al.*, 2011). Una ventaja asociada a la presencia de más de un sitio de unión es que la energía de Gibbs para un ligando multivalente es equiparable a la suma de las energías libres para cada uno de los sitios, lo que termodinámicamente es muy favorable (Weis y Drickamer, 1996).

Siempre ha existido controversia acerca de si existe cooperatividad en la unión de azúcares a lectinas que presentan multivalencia. La ausencia de cambios conformacionales al unirse las moléculas de ligando es consistente con un escenario de no cooperatividad. En 1978, Zentz y colaboradores proponían que para el caso de la ricina o existían dos sitios de unión independientes o dos sitios de igual afinidad pero con cooperatividad negativa (Zentz *et al.*, 1978). Sin embargo, algunos grupos ya especulaban con la posibilidad de una cooperatividad positiva entre los dos sitios de unión de la ricina (Adair y Kornfeld, 1974). Recientemente, se ha visto que a medida que la ricina va incorporando moléculas de lactosa, su estabilidad va aumentando. Por esta razón, se piensa que puede existir cierta cooperatividad que parece estar ligada a una pérdida de la flexibilidad que favorece de esta forma la acomodación de moléculas adicionales de ligando (Yao *et al.*, 2011).

### 3.5. Termodinámica de la interacción lectina-azúcar

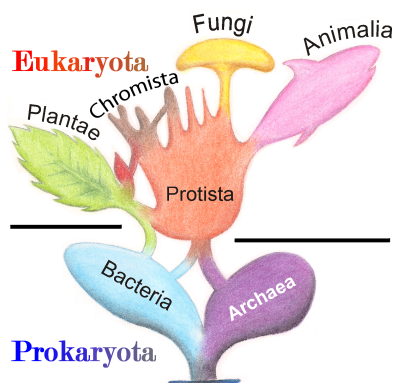
Además de la descripción de los contactos que se establecen entre los azúcares y los residuos del sitio de unión, una descripción completa de la interacción requiere una caracterización termodinámica, es decir, la determinación de la energía de Gibbs asociada al proceso de unión ( $\Delta G$ ), relacionada a su vez con los balances entálpico ( $\Delta H$ ) y entrópico ( $\Delta S$ ), vinculados a la interacción.

Los cambios de entalpía incluyen, por un lado, la contribución favorable de los contactos específicos proteína-carbohidrato y de los nuevos enlaces de hidrógeno generados entre moléculas de agua tras la desolvatación de la proteína y el carbohidrato y, por otro lado, la contribución desfavorable de la rotura de las interacciones proteína-agua y carbohidrato-agua. Los factores que afectan a la entropía de unión se relacionan con la variación de los grados de libertad translacional y rotacional del ligando y de la proteína como consecuencia de la formación del complejo, así como la reorganización del solvente durante este proceso. Aunque el sitio de unión pueda tener una organización predeterminada y la proteína sea, por tanto, capaz de unir el ligando sin cambios conformacionales importantes

en el sitio de unión, cualquier posible cambio en la flexibilidad y movilidad interna de la proteína también puede afectar a la termodinámica de la unión (López-Lucendo *et al.*, 2004). Como se deduce de los diferentes factores enumerados, las interacciones proteína-azúcar se caracterizan por una compensación entalpía-entropía que resulta en interacciones relativamente débiles en comparación con otra clase de interacciones, por ejemplo, las de tipo proteína-proteína. Así, se explica que las constantes de disociación ( $K_d$ ) para la mayoría de las interacciones lectina-monosacárido estén en el rango milimolar.

#### 4. Redescubriendo el reino de los hongos

En un principio, los botánicos y los microbiólogos fueron los primeros en sumergirse en el estudio de los hongos. Las contribuciones de ambos iniciaron el desarrollo posterior de la micología, una ciencia pluridisciplinar que estudia el reino de los hongos en todas sus facetas. El establecimiento de los hongos como un reino biológico independiente no fue una tarea fácil. Aunque inicialmente fueron incluidos en el reino de las plantas, con los avances en Biología Celular y Molecular, se consiguió dibujar un panorama taxonómico diferente (**Figura 1.12**) en el que los hongos aparecen englobados dentro del reino Fungi, junto con otros cuatro reinos más (Animalia, Chromista, Plantae y Protozoa) dentro del gran dominio Eukarya (Cavalier-Smith, 1998; Cavalier-Smith, 2004).



**Figura 1.12.** Representación del árbol filogenético de los seres vivos, según el sistema de clasificación actual, así como los principales procesos simbiogénéticos que dieron lugar al establecimiento de los diferentes reinos. Imagen cortesía de Maulucioni y Doridí ([www.wikipedia.com](http://www.wikipedia.com)).

Desde su descubrimiento, los hongos han sido de gran interés para el hombre. Desde un punto de vista científico, se han usado como modelo en Bioquímica y Genética Molecular ya que son fáciles de cultivar, se reproducen rápidamente y en grandes cantidades. En 1996 se completó la secuenciación del genoma de *Saccharomyces cerevisiae*, lo que contribuyó al estudio de la regulación génica, la caracterización de nuevas proteínas y la comprensión de mecanismos celulares comunes a otros organismos, como los animales (Goffeau *et al.*, 1996). Además, los hongos han sido y siguen siendo interesantes para la

industria biotecnológica, explotando procesos como la fermentación, tan útiles en la tecnología de alimentos, por ejemplo.

#### 4.1. Diversidad en el reino de los hongos

El número de especies fúngicas descubiertas hasta hoy sólo representa una pequeña fracción del total, que se estima en más de un millón y medio (Carlie *et al.*, 2001). A pesar de la gran cantidad, todas ellas presentan hifas (estructuras filamentosas largas y ramificadas constituidas por una o varias células y rodeadas de una pared celular) o al menos están estrechamente relacionadas con especies hifales. Además, en algún momento de su ciclo vital poseen una pared celular que contiene quitina (y no celulosa, como en las plantas) y se alimentan exclusivamente mediante nutrición absorptiva, es decir, absorben los nutrientes del entorno.

Teniendo en cuenta el gran número de especies fúngicas existentes, la clasificación de los hongos no es trivial y, de hecho, ésta va cambiando a medida que se va teniendo más información a nivel molecular. Las proteínas caracterizadas en este trabajo provienen de hongos que se engloban dentro del subreino Dikarya. Este subreino alberga dos filos diferentes: Ascomycota y Basidiomycota (Hibbett *et al.*, 2007). Estos filos incluyen especies que producen tanto hifas como células dicariotas, esto es, que contienen dos núcleos, y que, además, carecen de flagelos. El subreino Dikarya es también conocido como el de los hongos superiores, a pesar de que incluye especies anamorfas (con ciclo reproductivo asexual) como los mohos. Filogenéticamente, los dos filos componentes del subreino constituyen un mismo grupo evolutivo (Lutzoni *et al.*, 2004; James *et al.*, 2006).

##### 4.1.1. Ascomycota

El filo Ascomycota es el más numeroso del reino Fungi, llegando a contabilizarse cerca de 64.000 especies. Los miembros de este filo son aquellos hongos cuyo ciclo de reproducción sexual requiere de la producción de ascosporas haploides a través de la meiosis de núcleos diploides en un asca. Muchos hongos ascomicetos producen sus ascas en complejos cuerpos fructíferos<sup>9</sup> llamados ascocarpos. Estos hongos suelen llamarse con el nombre de ascomicetos verdaderos y se clasifican por la forma del ascocarpo. Otros hongos dentro de este grupo también llevan a cabo una reproducción de tipo asexual mediante esporulación, produciendo conidios en los conidióforos, estructuras especializadas en la producción de conidios localizadas en el extremo de las hifas.

#### 4.1.2. Basidiomycota

El filo Basidiomycota es el más evolucionado y uno de los más variados y reconocidos, pues entre sus cerca de 22.000 especies se encuentran las clásicas setas. Sus miembros, conocidos también como basidiomicetos, son hongos filamentosos compuestos por hifas (a excepción de las levaduras) y se caracterizan porque durante la reproducción sexual producen basidiosporas haploides provenientes de un basidio en el cual núcleos diploides sufren meiosis. Los basidios se localizan en los cuerpos fructíferos (basidiocarpos), que representan tan solo un 10% del hongo, ya que el resto, el micelio, suele estar bajo tierra. Cabe señalar que una minoría de hongos basidiomicetos puede también reproducirse de forma asexual.

#### 4.2. Interacción de los hongos con el entorno

Los hongos establecen diferentes tipos de relaciones con el resto de organismos que se encuentran en su mismo entorno. Desde el punto de vista de las interacciones con el medio, unos de los hongos más importantes son los hongos micorrícicos, constituidos por un grupo heterogéneo de hongos terrestres. Estos hongos se hallan en muchos ecosistemas donde establecen asociaciones simbióticas con las raíces mediante lo que se denomina una micorriza<sup>25</sup>. De hecho, se estima que entre el 90 y el 95% de las plantas terrestres presentan micorrizas de forma habitual (Wang y Qiu, 2006). Al igual que en otras relaciones simbióticas, en las micorrizas ambos participantes obtienen beneficios. Las plantas suelen recibir del hongo agua y nutrientes minerales, por ejemplo fósforo, mientras que el hongo obtiene de la planta hidratos de carbono y vitaminas que por sí mismo es incapaz de sintetizar (Harrison, 2005). También se ha visto que las micorrizas suponen una ventaja defensiva para la planta, ya que los hongos las protegen de agentes patógenos, como otros hongos nocivos y nemátodos<sup>29</sup> (Smith y Read, 2008).

Según su morfología, las micorrizas se dividen en dos grandes grupos: ectomicorrizas y endomicorrizas. En las ectomicorrizas, el hongo rodea a la raíz, formando el manto, y, por tanto, las hifas no penetran en el interior de las células de la planta. Este tipo de micorrización es típica entre los árboles de zonas templadas, como hayas, robles y pinos. Tanto hongos ascomicetos como basidiomicetos llevan a cabo este tipo de simbiosis. Por otro lado, en las endomicorrizas, el hongo penetra en la raíz, formándose vesículas alimenticias y arbusculos. Suelen darse en todo tipo de plantas, pero predominan en las hierbas y gramíneas.

Algunos hongos no establecen relaciones simbióticas con las plantas, sino que actúan como patógenos de éstas. Los hongos fitopatógenos emplean una gran variedad de estrategias de infección que les permiten colonizar las plantas. Existen especies de hongos que son patógenos especie-dependientes, mientras que otros tienen menos restricciones a la hora de actuar. Generalmente, estos hongos se clasifican en tres grupos: (a) hongos necrótrofos, que matan a las células de su alrededor y se alimentan del tejido muerto; (b) hongos biótrophos, que crecen dentro de la planta y toman nutrientes de ella; y (c) hemibiótrofos, capaces de emplear ambas formas de alimentación (Amselem *et al.*, 2011). Los mecanismos intrínsecos a la infección son complejos y todavía no están bien esclarecidos. Por ejemplo, en el caso de los hongos con alimentación necrotrófica, se sabe que son capaces de secretar proteínas efectoras que son internalizadas por las células del huésped, interactuando con el proceso de expresión génica, lo que inicia de algún modo la enfermedad (Oliver y Solomon, 2010).

El medio natural es una fuente de recursos, pero también fuente de retos hostiles, donde los hongos tienen que convivir con parásitos, patógenos y depredadores. Dado que los hongos no poseen un sistema inmunológico, necesitan desarrollar mecanismos de defensa frente a los potenciales ataques que puedan surgir. Teniendo en cuenta que la mayoría de hongos, al igual que las plantas, no presentan motilidad, se puede desprender la idea de que usen mecanismos de defensa similares a éstas. De hecho, una gran cantidad de compuestos químicos, resultado del metabolismo secundario, han demostrado ser una potente amenaza frente a organismos micófilos (Rohlf *et al.*, 2007; Fox y Howlett, 2008). Aún así, se ha observado que las proteínas son las máximas responsables de las actividades insecticidas en las setas y no otro tipo de compuestos químicos (Wang *et al.*, 2002).

## 5. Lectinas fúngicas

### 5.1. Distribución, bioquímica, funciones y estructura

La primera descripción de una lectina fúngica data de 1891 cuando Kobert descubrió un agente venenoso en la seta *Amanita phalloides*. Más tarde se vio que este agente se trataba realmente de una lectina hemolítica (Horejsi y Kocourek, 1978). En los últimos años, las lectinas fúngicas han atraído la atención de la comunidad científica por su interés en medicina, como agentes antitumorales, antiproliferantes o inmunomoduladores. Muchas de estas lectinas están empezando a ser purificadas con el fin de caracterizarlas funcional y es-



tructuralmente en espera de dilucidar el papel biológico que desempeñan en la naturaleza y explorar el desarrollo de potenciales aplicaciones biotecnológicas novedosas.

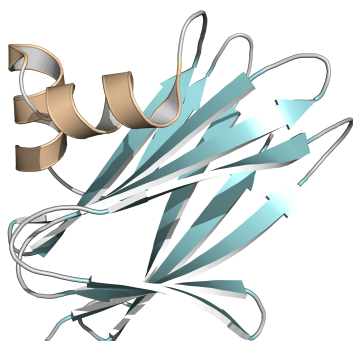
Las lectinas de hongos se han aislado de diferentes partes, incluyendo el micelio, conidios, esporocarpos y cuerpos fructíferos. En algunos hongos se ha documentado la producción de lectinas extracelulares, como es el caso de *Macrophomina phaseolina* (Bhowal *et al.*, 2005). Generalmente, estas lectinas son identificadas mediante ensayos de aglutinación celular o por análisis de similitud de secuencia, aislándolas directamente del hongo o produciéndolas de forma recombinante.

Hasta el momento, se ha visto que existe una amplia variedad de lectinas desde un punto de vista bioquímico. Algunas presentan glicosilaciones, llegando a representar más del 30% de la masa total, como en la lectina de *Rigidoporus lignosus* (Guillot y Konska, 1997), o muestran una ausencia total de modificaciones, tal y como se observa en la lectina de *Laccaria amethystea* (Guillot *et al.*, 1983). En general, no necesitan metales para llevar a cabo su actividad, aunque se ha observado que en algunos casos, la presencia de estos puede aumentar la actividad hemaglutinante, como en la lectina de *Xerocomus chrysenteron* (Liu *et al.*, 2004), o inhibirla, como en la lectina de *Inocybe umbrinella* (Zhao *et al.*, 2009). Es destacable que la mayoría de las lectinas de hongos son estables en un amplio rango de pH y altas temperaturas, observándose, como media, temperaturas de desnaturalización por encima de 50 °C (Khan y Khan, 2011).

Aunque todavía queda mucho por investigar sobre las funciones biológicas de las lectinas fúngicas, ya se han empezado a comprender algunos de los procesos en los que están involucradas. La habilidad de las lectinas para reconocer azúcares de diferente complejidad les otorga la capacidad de participar en gran variedad de procesos fisiológicos. Una de las primeras funciones que se observó estaba relacionada con el crecimiento y la morfogénesis de los hongos. En efecto, se vio cómo la lectina de *Pleurotus cornucopiae* está implicada en la cohesión de las hifas durante el desarrollo de los basidiocarpos (Kaneko *et al.*, 1993). Más tarde, se observó que las galectinas CGL1 y CGL2 de *Coprinopsis cinerea* participan en el desarrollo del himenio (Walser *et al.*, 2005); mientras que la lectina de *Paracoccidioides brasiliensis* interacciona con la quitina de la pared celular, influyendo directamente en el crecimiento del hongo (Ganiko *et al.*, 2007). Otras lectinas están directamente involucradas en las simbiosis que se establecen entre hongos y árboles o bacterias (Giollant *et al.*, 1993; Cerigini *et al.*, 2008). La toxicidad de muchos cuerpos fructíferos también se asocia a lectinas citosólicas, que parecen estar implicadas en los mecanismos de defensa que poseen los hongos contra parásitos y depredadores. Como ejemplo, cuando los orga-

nismos modelo *Caenorhabditis elegans* y *Drosophila melanogaster* crecen en presencia de la lectina CCL2 de *C. cinerea* su supervivencia se ve mermada, lo que sugiere que la lectina actúa como una toxina contra ellos (Schubert *et al.*, 2012). Finalmente, se ha visto que muchas lectinas fúngicas están involucradas en patogénesis humana. En este sentido, por ejemplo, la lectina AFL1 de *Aspergillus fumigatus*, hongo responsable de graves infecciones nosocomiales en pacientes inmunodeprimidos, presenta gran afinidad frente a oligosacáridos de la sangre humana ricos en fucosilaciones y produce la inflamación de las células bronquiales (Houser *et al.*, 2012).

Con respecto a las estructuras de las lectinas fúngicas que ya se disponen, se observa cómo muchas muestran plegamientos que ya habían sido previamente descritos para lectinas de otros tipos de organismos. Por ejemplo, la lectina de *Aleuria aurantia*, cuya estructura fue la primera en resolverse para una lectina de un hongo, presenta un plegamiento de hélice enrollada con seis láminas  $\beta$  dispuestas alrededor de un eje central (Wimmerova *et al.*, 2003). Otras lectinas, como la aglutinina de *Marasmius oreades* (MOA), adquieren una estructura de tipo trébol  $\beta$  con tres sitios operativos de unión a azúcares (Grahn *et al.*, 2007; Grahn *et al.*, 2009). También se han descubierto lectinas con plegamiento tipo galectina, como la lectina dimérica de *Agrocybe aegerita* (Yang *et al.*, 2005; Yang *et al.*, 2009). No obstante, se ha identificado lo que parece ser una nueva familia estructural de lectinas que curiosamente hasta el momento solo se ha visto en hongos. La estructura consiste en dos láminas  $\beta$ , de cuatro y seis hebras  $\beta$  cada una, conectadas por dos hélices  $\alpha$  (**Figura 1.13**). Por ahora, dentro de esta familia se han identificado las lectinas de *Agaricus bisporus* (Carrizo *et al.*, 2005), *Sclerotium rolfii* (Leonidas *et al.*, 2003) y *Xerocomus chrysantheron* (Birck *et al.*, 2004).



**Figura 1.13.** Representación de bucles y cintas de la estructura terciaria de la lectina de *Sclerotium rolfii*. El plegamiento de esta lectina es novedoso y sólo se ha identificado en lectinas fúngicas.

La especificidad mostrada por las lectinas fúngicas es también muy variable y contempla desde azúcares sencillos a complejas estructuras localizadas en glicoproteínas (**Tabla 1.1**). En general, presentan constantes de baja afinidad para mono y oligosacáridos,

aunque la afinidad aumenta cuanto más complejo es el azúcar que reconocen, ya que, por ejemplo, la lectina de *Fusarium solani* tiene una constante de disociación por asialomucina del orden de picomolar (Khan *et al.*, 2007). En la mayoría de los casos, la unión es exotérmica y está gobernada por los cambios entálpicos.

Lectina	Organismo	Especificidad	Referencia
APL	<i>Amanita pantherina</i>	Mucina	(Zhuang <i>et al.</i> , 1996)
CNL	<i>Clitocybe nebularis</i>	N,N'-diacetillactosadamina	(Pohleven <i>et al.</i> , 2012)
FSL	<i>Fusarium solani</i>	Asialomucina	(Khan <i>et al.</i> , 2007)
LAF	<i>Laccaria amethystina</i>	L-fucosa	(Guillot <i>et al.</i> , 1983)
LDL	<i>Lactarius deliciosus</i>	Gal $\beta$ (1 $\rightarrow$ 3)GalNAc	(Guillot <i>et al.</i> , 1991)
MPL	<i>Macrophomina phaseolina</i>	N-acetillactosamina	(Bhowal <i>et al.</i> , 2005)
PSL	<i>Peziza sylvestris</i>	Arabinosa	(Wang y Ng, 2005)
SCL	<i>Schizophyllum commune</i>	N-acetil-D-galactosamina	(Chumkhunthod <i>et al.</i> , 2006)

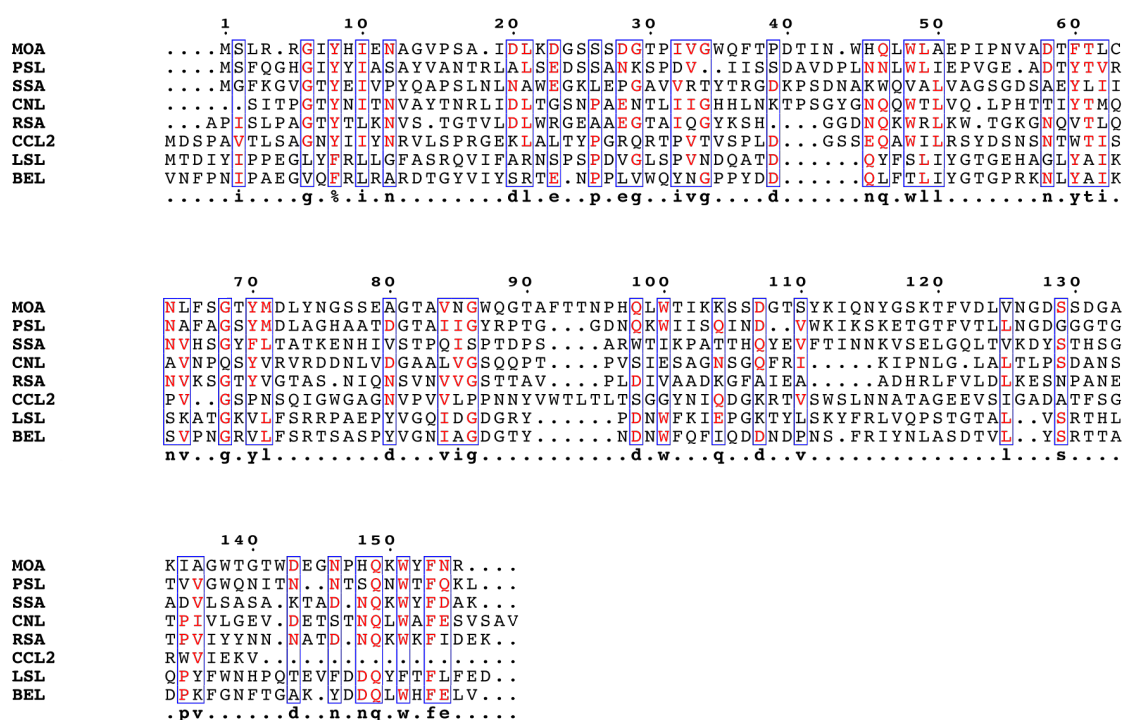
**Tabla 1.1.** Especificidad de algunas lectinas fúngicas

## 5.2. Lectinas fúngicas con estructura trébol $\beta$

Hasta la fecha se han descrito las estructuras de ocho lectinas fúngicas con estructura trébol  $\beta$ : la aglutinina de *M. oreades* (MOA) (Grahm *et al.*, 2007), la lectina de *Laetiporus sulphureus* (LSL) (Mancheño *et al.*, 2005), la aglutinina de *Sclerotinia sclerotiorum* (SSA) (Sulzenbacher *et al.*, 2010), la lectina de *Polyporus squamosus* (PSL) (Kadirvelraj *et al.*, 2011), la lectina de *Clitocybe nebularis* (CNL) (Pohleven *et al.*, 2012), la lectina de *C. cinerea* (CCL2) (Schubert *et al.*, 2012), la aglutinina de *Rhizoctonia solani* (RSA) (Skamnaki *et al.*, 2013) y la lectina de *Boletus edulis* (BEL) (Bovi *et al.*, 2013). Todas ellas son originarias de hongos basidiomicetos, a excepción de SSA, que procede de un hongo ascomiceto. Además, la mayoría se expresan sobre todo en los cuerpos fructíferos, menos SSA y RSA que se almacenan en el esclerocio o micelio. Las lectinas PSL, CNL y CCL2 fueron expresadas de forma recombinante en *E. coli*, mientras que el resto se aislaron directamente del hongo.

Como se había descrito anteriormente para el caso general de las proteínas con estructura trébol  $\beta$ , estas lectinas fúngicas manifiestan baja identidad de secuencia a pesar de que adoptan el mismo plegamiento (**Figura 1.14**). No obstante, en todas ellas se identifica el motivo (QXW)<sub>3</sub> clásico de los tréboles  $\beta$ . En ningún caso se llega a un porcentaje de identidad de secuencia entre ellas superior al 50%, siendo la media de éste próxima a un 16% (**Tabla 1.2A**). Sin embargo, como cabría esperar, las desviaciones cuadráticas medias o

RMSD (del inglés *Root-Mean-Square Deviation*) que se obtienen al comparar las estructuras tridimensionales son relativamente bajas. La comparación entre LSL y MOA da lugar al valor de RMSD más elevado, 2,45 Å; mientras que el alineamiento estructural de CNL y RSA presenta un RMSD de solo 1,11 Å. El valor del RMSD medio que se obtiene al comparar todas las estructuras es de 1,78 Å, lo que indica cómo de bien se conserva la estructura de estas lectinas en las diferentes especies (**Tabla 1.2B**).



**Figura 1.14.** Alineamiento de las secuencias de aminoácidos de las ocho lectinas fúngicas que presentan un plegamiento trébol  $\beta$  (de las lectinas modulares MOA, LSL y PSL sólo se ha utilizado la secuencia correspondiente a los respectivos módulos lectina N-terminales). Los residuos conservados en más del 50% de las secuencias aparecen en rojo y recuadrados en azul. En la parte de abajo del alineamiento y en *negrita* se muestran los residuos cuyo consenso es mayor del 50%. El alineamiento se ha realizado utilizando el programa Multalin (Corpet, 1988).

En relación con la estructura cuaternaria, no todas estas lectinas adquieren el mismo estado de oligomerización (**Tabla 1.3**). Tan solo las lectinas CCL2 y CNL son monoméricas, coincidiendo que son asimismo dos de las más cortas en secuencia (142 y 149 aminoácidos, respectivamente) entre las lectinas procedentes de cuerpos fructíferos. Sin embargo, SSA, RSA y BEL, que también tienen secuencias pequeñas (153, 137 y 146 aminoácidos, respectivamente), son proteínas homodiméricas, lo que muestra un mecanismo novedoso de dimerización en esta clase de proteínas. El resto de las lectinas, MOA, LSL y PSL, son proteínas modulares (tienen secuencias más extensas) y oligoméricas. En el extremo N-terminal (de ~150 residuos) contienen un módulo lectina con el plegamiento tipo cade-

na B de la ricina, mientras que en el extremo C-terminal presentan un módulo responsable de la oligomerización y con alguna potencial actividad catalítica que determina su función biológica. Por ejemplo, en el caso de MOA, su extremo C-terminal adopta un plegamiento tipo  $\alpha/\beta$  en donde se pueden identificar residuos de una potencial tríada catalítica (Grahm *et al.*, 2007). Los estados oligoméricos para las tres lectinas son diferentes: MOA es un homodímero, en el caso de PSL se establece un equilibrio dímero-tetrámero y LSL se comporta como un hexámero en solución, tal y como muestran los resultados de los ensayos de ultracentrifugación analítica (Mancheño *et al.*, 2005).

**A**

Porcentaje de identidad de secuencia (%)

	MOA	LSL	SSA	PSL	CNL	CCL2	RSA	BEL
MOA		10	11	40	13	12	26	11
LSL	10		8	13	11	7	13	41
SSA	11	8		13	15	11	18	8
PSL	40	13	13		19	15	25	13
CNL	13	11	15	19		13	29	11
CCL2	12	7	11	15	13		16	6
RSA	26	13	18	25	29	16		13
BEL	11	41	8	13	11	6	13	

**B**Desviación cuadrática media (RMSD) con respecto a los C $\alpha$  (Å)

	MOA	LSL	SSA	PSL	CNL	CCL2	RSA	BEL
MOA		2,45	1,44	1,32	1,36	1,71	1,24	1,97
LSL	2,45		2,16	2,19	2,24	2,38	2,03	1,34
SSA	1,44	2,16		1,42	1,69	1,91	1,23	1,97
PSL	1,32	2,19	1,42		1,73	1,78	1,39	2,04
CNL	1,36	2,24	1,69	1,73		1,77	1,11	1,88
CCL2	1,71	2,38	1,91	1,78	1,77		1,68	2,45
RSA	1,24	2,03	1,23	1,39	1,11	1,68		2,01
BEL	1,97	1,34	1,97	2,04	1,88	2,45	2,01	

**Tabla 1.2.** A. Porcentajes de identidad de secuencia entre las lectinas fúngicas cuya estructura presenta un plegamiento trébol  $\beta$ . Los porcentajes se han determinado a partir del programa ClustalW2 (EMBL-EBI). B. Valores de las desviaciones cuadráticas medias (RMSD) con respecto a los C $\alpha$  entre las estructuras de las diferentes lectinas fúngicas. Los RMSD se han determinado utilizando el servidor PDBeFold (EMBL-EBI).

Todas estas lectinas muestran sitios operativos de unión de azúcares (**Tabla 1.3**). A diferencia del resto, MOA y BEL son las únicas de ellas capaces de unirlos en los tres subsitios ( $\alpha$ ,  $\beta$  y  $\gamma$ ). LSL y RSA, en cambio, tienen dos sitios operativos (los sitios  $\beta$  y  $\gamma$  y los sitios

$\alpha$  y  $\gamma$ , respectivamente). El resto, CCL2, CNL, SSA y PSL, solo poseen un único sitio. Las interacciones que se observan entre la lectina y el azúcar no son iguales en todos los casos, aunque siempre prevalecen los enlaces de hidrógeno junto con el apilamiento de algún residuo aromático y la participación de moléculas de agua. Por ejemplo, a pesar de que en PSL el sitio operativo de unión es el  $\beta$ , hay residuos del sitio  $\alpha$  implicados también en interacciones con el ligando. Por otro lado, en CCL2 se ha descrito un modo completamente nuevo de unión, ya que el ligando queda adherido en la interfase entre los sitios  $\beta$  y  $\gamma$  mediante numerosos enlaces de hidrógeno.

Lectina	Estado oligomérico	Sitios operativos	Especificidad
MOA	Dímero	$\alpha$ , $\beta$ y $\gamma$	Gal $\alpha$ (1 $\rightarrow$ 3)Gal
LSL	Hexámero	$\beta$ y $\gamma$	Lactosa y LacNAc
SSA	Dímero	$\alpha$	Gal y GalNAc
PSL	Dímero $\rightleftharpoons$ Tetrámero	$\beta$ + $\alpha$	Neu5Ac $\alpha$ (2 $\rightarrow$ 6)Gal
CNL	Monómero	$\alpha$	LacdiNAc
CCL2	Monómero	Interfase $\beta$ - $\gamma$	Fuc $\alpha$ (1 $\rightarrow$ 3)GlcNAc
RSA	Dímero	$\alpha$ y $\gamma$	GalNAc
BEL	Dímero	$\alpha$ , $\beta$ y $\gamma$	$\beta$ -galactósidos

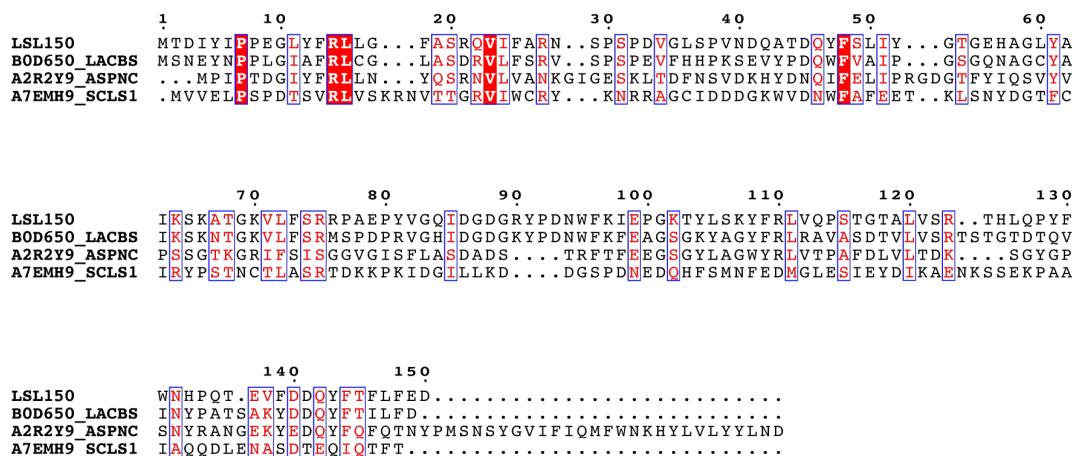
**Tabla 1.3.** Lista de algunas de las propiedades bioquímicas de las lectinas fúngicas con estructura trébol  $\beta$ .

Excepto para las proteínas LSL y BEL, se han llevado a cabo diferentes ensayos para averiguar la especificidad de unión de las lectinas. En todos los casos se pone de manifiesto la preferencia por los residuos de galactosa sobre cualquier otro tipo de monosacárido (**Tabla 1.3**). La única excepción la representa la proteína CCL2, que presenta mayor afinidad por carbohidratos que contienen el motivo Fuc $\alpha$ (1 $\rightarrow$ 3)GlcNAc, lo que puede deberse al distintivo modo de unión descrito. Generalmente, todos los ligandos característicos para estas lectinas están N-acetilados. Incluso en el caso de LSL, la estructura está resuelta en complejo con dos moléculas de N-acetilactosamina (LacNAc), viéndose como se establecen contactos adicionales, que podrían ser responsables del incremento en la afinidad. De hecho, ensayos de hemaglutinación y hemólisis en presencia de diferentes inhibidores demostraron que LacNAc es uno de los candidatos más potentes (Tateno y Goldstein, 2003).

En términos generales, se puede afirmar que actualmente no disponemos de información suficiente acerca de la función biológica de estas lectinas fúngicas. En el caso de las lectinas CCL2 y CNL, éstas parecen participar en los mecanismos de defensa de los hongos frente a organismos parásitos y depredadores invertebrados, como insectos y nemátodos.

De hecho, los residuos de fucosa en posición  $\alpha(1\rightarrow3)$  que reconoce específicamente la lectina CCL2 son comunes en las estructuras de glicoproteínas del nemátodo *C. elegans*. De esta manera, cuando es alimentado con CCL2 marcada con una sonda fluorescente, se observa cómo la proteína se localiza exclusivamente en el intestino del nemátodo. En el resto de las lectinas, la función biológica se desconoce, aunque se cree que deben de estar involucradas en funciones de almacenamiento, como mediadores de la morfogénesis o participando en los mecanismos de defensa.

En este contexto, el presente trabajo trata de profundizar en el conocimiento de las lectinas fúngicas con estructura trébol  $\beta$  y en entender los determinantes estructurales de la interacción lectina-azúcar. Tomando como referencia la lectina del hongo *Laetiporus sulphureus*, se han descrito tres proteínas fúngicas que muestran un porcentaje de identidad de secuencia significativo con el módulo N-terminal de LSL: la proteína B0D650\_LACBS de *Laccaria bicolor* (Martin *et al.*, 2008), con un 52% de identidad; la proteína A7EMH9\_SCLS1 de *Sclerotinia sclerotiorum* (Amselem *et al.*, 2011), con un 34%; y la proteína A2R2Y9\_ASPNC de *Aspergillus niger* (Pel *et al.*, 2007), con un 28% (**Figura 1.15**). Aunque es evidente que existen numerosos homólogos estructurales del dominio lectina N-terminal de LSL, como se ha dicho con antelación acerca de la conservación de la estructura de los tréboles  $\beta$ , estos muestran similitudes de secuencia poco significativos por lo que nuestro interés se ha centrado en las cuatro proteínas anteriores, incluyendo la lectina *L. sulphureus*.



**Figura 1.15.** Alineamiento de las secuencias de los homólogos cercanos al módulo lectina N-terminal de LSL (LSL<sub>150</sub>). Los residuos conservados en todas las proteínas se muestran sobre fondo rojo. Los residuos conservados en más del 50% de las secuencias aparecen en rojo y recuadrados en azul. El alineamiento se ha realizado utilizando el programa Multalin (Corpet, 1988).



### 5.2.1. La lectina LSL de *Laetiporus sulphureus*

La lectina de *Laetiporus sulphureus* (LSL, de *Laetiporus sulphureus* *Lectin*) fue aislada inicialmente a partir de los carpóforos o basidiocarpos de este hongo parasítico (Konska *et al.*, 1994). *L. sulphureus* o seta de pollo (**Figura 1.16**) es un hongo basidiomiceto, perteneciente al orden Polyporales, que crece sobre árboles vivos y heridos, y de forma excepcional sobre árboles muertos. Generalmente, aparece asociado a robles y castaños. Los cuerpos fructíferos aparecen una vez al año en cualquier época, menos en invierno, llegando a proyectarse hasta 20 cm del tronco del árbol y no presentando un espesor mayor de 2 cm. De esta especie destaca sobre todo el color, de un anaranjado intenso y llamativo.



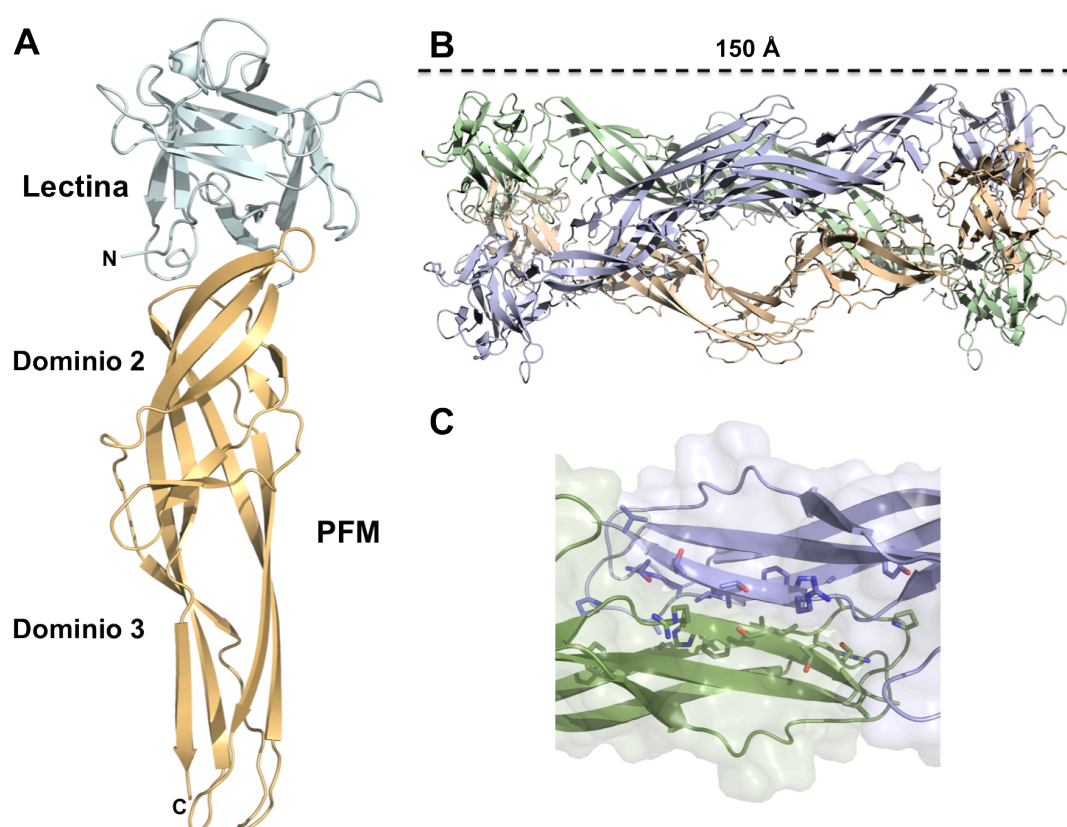
**Figura 1.16.** Imagen del hongo basidiomiceto y parásito *Laetiporus sulphureus*. Esta especie se localiza fundamentalmente en los bosques de América del Norte, aunque también se han encontrado ejemplares en Europa. Fotografía cortesía de Michael Kuo ([www.mushroomexpert.com](http://www.mushroomexpert.com)).

La lectina LSL se purificó por primera vez mediante cromatografía de afinidad a Sepharose® 4B (constituido por unidades de D-galactosa y 3,6-anhidro-L-galactosa unidas mediante enlaces  $\alpha(1\rightarrow3)$  y  $\beta(1\rightarrow4)$ ). La elución se llevó a cabo con lactosa, demostrándose, en consecuencia, la capacidad que tiene la proteína para unir azúcares. Mediante diversos ensayos con eritrocitos humanos del grupo sanguíneo O, se observó que la proteína demostraba tener, al igual que muchas otras lectinas, actividad hemaglutinante y hemolítica. La inhibición de estas actividades en presencia de diferentes azúcares indicó que la lectina muestra cierta especificidad por azúcares que contienen motivos de LacNAc en su estructura (Konska *et al.*, 1994).

La posterior clonación de LSL a partir del ADN codificante (ADNc) de los cuerpos fructíferos permitió describir la presencia de tres isoformas: LSLa, LSLb y LSLc. Estas isoformas o isolectinas comparten entre un 80-90% de identidad de secuencia y muestran similitud estructural con toxinas bacterianas, como la toxina  $\alpha$  de *Clostridium septicum* (Tateno y Goldstein, 2003). La comparación de las tres secuencias permitió deducir que LSLa es realmente la isoforma que se había caracterizado previamente a partir de la extracción directa del hongo. De las tres, solo LSLa y LSLb pudieron purificarse mediante



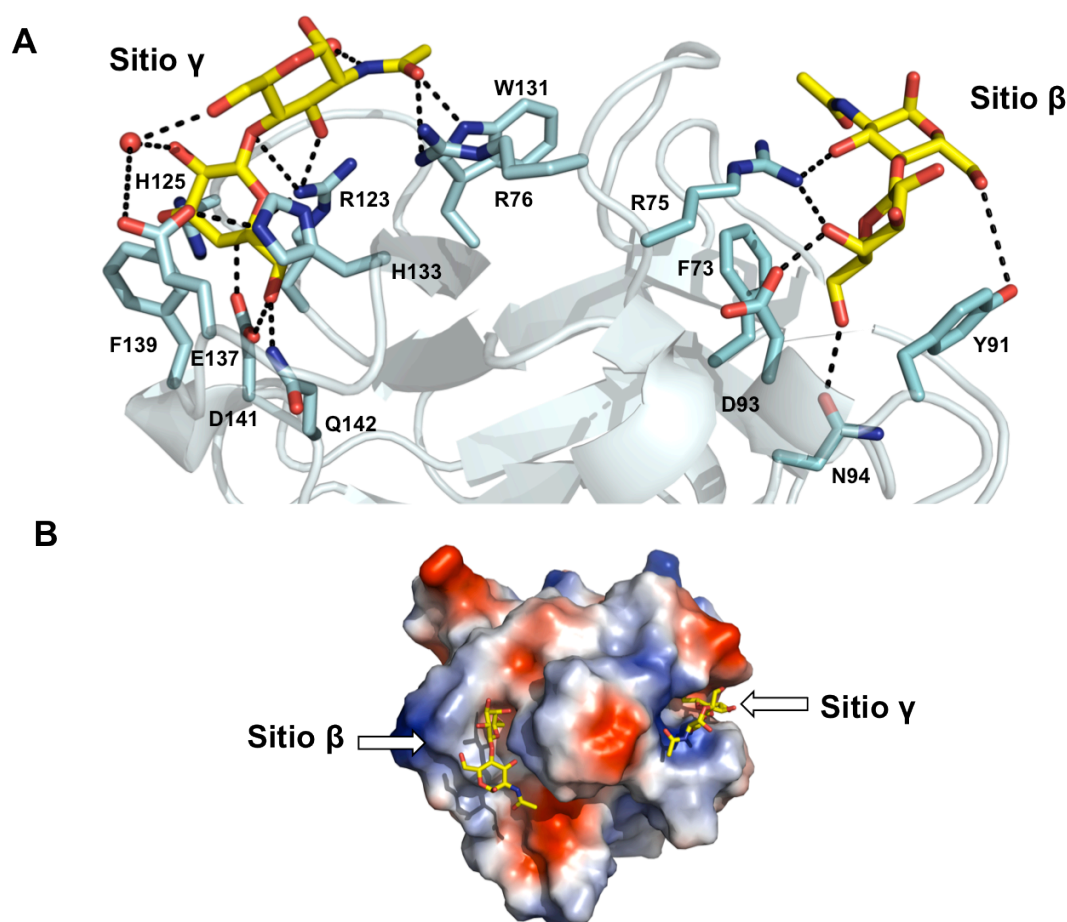
cromatografía de afinidad ya que LSLc agregaba formando cuerpos de inclusión. Las diferencias observadas en sus secuencias, sobre todo en los aminoácidos del entorno de los sitios de unión, probablemente sean las responsables de que LSLa y LSLb no se comporten igual en los ensayos de inhibición de las actividades hemaglutinante y hemolítica. Mientras que LSLa se ve fuertemente inhibida por LacNAc, LSLb no muestra la misma tendencia. Cabe destacar que LSLb no conserva la Phe-139, responsable, como se verá más adelante, del apilamiento con el anillo de galactosa en el sitio  $\gamma$ .



**Figura 1.17.** Estructura tridimensional de la lectina LSLa de *Laetiporus sulphureus*. **A.** Modelo de *bucles y cintas* donde se muestra el extremo lectina N-terminal (azul) y el módulo C-terminal formador de poros (PFM) (naranja). **B.** Estructura hexamérica de LSLa. **C.** Vista ampliada de la zona de contacto entre dos monómeros de LSLa a través del dominio 3 del PFM. La asociación permite el establecimiento de una extensa lámina  $\beta$  antiparalela formada a partir de los sándwiches  $\beta$  de cada protómero.

La resolución estructural de la lectina LSLa fue publicada en el año 2005 (Mancheño *et al.*, 2005). La estructura se obtuvo mediante cristalografía de rayos X a 2,6 Å de resolución, aplicando reemplazo isomorfo con dispersión anómala (SIRAS). La estructura reveló la presencia de dos módulos bien diferenciados (**Figura 1.17A**). El extremo N-terminal (primeros 150 residuos) se corresponde con el módulo lectina con plegamiento trébol  $\beta$ , que posee dos sitios operativos de unión a azúcares (sitios  $\beta$  y  $\gamma$ ) tal y como se desprende de las estructuras de los complejos con lactosa y LacNAc. Por otro lado, el módulo C-

terminal (residuos 151-315) exhibe una estructura tridimensional constituida por dos dominios (dominios 2 y 3) que se asemejan a los dominios 3 y 4 de la toxina formadora de poros aerolisina de la bacteria Gram negativa *Aeromonas hydrophila*, y a los dominios 2 y 3 de la toxina  $\epsilon$  de *Clostridium perfringens*.



**Figura 1.18.** Unión de N-acetilglucosamina (LacNAc) al módulo lectina de LSLa. **A.** Estructura tridimensional de los sitios de unión  $\beta$  y  $\gamma$  con sendas moléculas de LacNAc unidas (representadas en *amarillo* y con el modelo de *varillas*). Se muestran los residuos de los sitios de unión que interactúan con los ligandos, así como las principales interacciones que se establecen. **B.** Superficie molecular del trébol  $\beta$  con dos moléculas de LacNAc unidas. La superficie está coloreada de acuerdo al potencial electrostático, *azul* para la carga positiva y *rojo* para la negativa.

Las interacciones que se establecen entre los azúcares y la lectina comprenden enlaces de hidrógeno, directos o mediados por moléculas de agua, e interacciones de van der Waals (**Figura 1.18**). El sitio  $\beta$  contribuye con la participación de los residuos Tyr-91, que mediante una interacción C-H/ $\pi$  se ancla a la cara B del anillo de galactosa, y Asp-93 y Asn-94, que establecen enlaces de hidrógeno con los oxígenos O4 y O6 de la galactosa. En el caso del ligando LacNAc, la voluminosa cadena lateral de la Phe-73, que interacciona a su vez con la Arg-75 mediante interacciones catión- $\pi$ <sup>21</sup>, contribuye a la desestabilización de la

unión porque impide el establecimiento de moléculas de agua que estabilicen el grupo acetamido, como, sin embargo, ocurre en el otro sitio de unión. Por otro lado, en el sitio  $\gamma$  se establece un apilamiento con el anillo del residuo Phe-139, mientras que los principales enlaces de hidrógeno los lleva a cabo con los residuos Asp-141 y Gln-142. Los residuos Arg-123, His-125 e His-133 contribuyen asimismo con interacciones que implican al oxígeno glicosídico y al oxígeno del anillo de galactosa, así como a los oxígenos O4 y O4' de los anillos de galactosa y glucosa, respectivamente. De esta forma, el mayor número de interacciones justificaría la mayor afinidad por el sitio  $\gamma$ .

El extremo C-terminal de la lectina LSLa se corresponde con el módulo formador de poros (PFM, por sus siglas en inglés *Pore-Forming Module*). Es, por tanto, el responsable de la actividad hemolítica. Posiblemente, el módulo lectina actuaría como vehículo, dirigiendo al PFM para que pudiese llevar a cabo su función específicamente sobre su diana biológica. PFM muestra una estructura todo  $\beta$  con once hebras  $\beta$  y una pequeña hélice  $3_{10}$ . El dominio 2 está constituido por una lámina de cinco hebras  $\beta$  antiparalelas cubierta en uno de los lados por un bucle o *loop* anfipático, y el dominio 3 posee un plegamiento tipo sándwich  $\beta$ . Este último dominio está, a su vez, directamente implicado en la formación del hexámero estable en solución (**Figura 1.17B**), que da lugar a la formación de una amplia lámina  $\beta$  antiparalela (**Figura 1.17C**). Aunque el área de contacto es pequeña, es altamente hidrofóbica (los residuos Ala-252, Phe-254, Val-256, Ala-258, Leu-306 y Leu-310 son los que están directamente implicados en esta región de contacto), lo que favorece la existencia de fuertes interacciones de tipo apolar. El carácter oligomérico en solución descrito plantea numerosos interrogantes acerca del mecanismo de reorganización que debe experimentar la lectina para llevar a cabo la formación de poros sobre las membranas. La formación de poros por LSLa sería similar al mecanismo descrito para los miembros de la familia de la aerolisina, como la toxina  $\alpha$  de *C. septicum* (Melton *et al.*, 2004): se iniciaría con la formación de un preporo oligomérico en el plano de la membrana seguido de la formación del poro funcional mediante la inserción del *loop* anfipático en la bicapa lipídica.

### 5.2.2. La lectina LBL de *Laccaria bicolor*

*Laccaria bicolor*, también conocida como seta de dos colores, es un hongo basidiomiceto perteneciente al orden de los Agaricales (**Figura 1.19**). Se localiza en zonas templadas junto a pinos y abedules, fructificando a finales de verano y otoño. Presenta un sombrero que puede llegar a medir entre 2-5 cm de diámetro en ejemplares adultos con forma

convexa o plano-convexa. Su característica principal es su color ocre parduzco con matices rojizos. El pie es alargado y esbelto y adquiere coloración violácea en la base, donde comienza el micelio.



**Figura 1.19.** Cuerpos fructíferos del hongo basidiomiceto y ectomicorrícico *Laccaria bicolor*. El hongo fructifica en las zonas templadas y bosques boreales de América del Norte y Europa. Fotografía cortesía de D. Vairelles.

La secuencia del genoma completo de *L. bicolor* se reportó en el año 2008 (Martin *et al.*, 2008). Esto supuso que *L. bicolor* se convirtiese en el primer hongo simbiótico ectomicorrícico cuyo genoma fue secuenciado. Debido a la importancia de la micorrización en el desarrollo de plantas y hongos, el genoma de *L. bicolor* ha establecido las bases para estudios genéticos y bioquímicos dirigidos a esclarecer los mecanismos de simbiosis (Martin y Selosse, 2008). Su interés es mayor, ya que *L. bicolor* se ha descrito como un organismo beneficioso para el desarrollo de ciertas especies arbóreas en la reforestación de bosques por su capacidad de micorrización (Villeneuve *et al.*, 1991). El genoma tiene una extensión de 65 megabases (Mpb) y alberga unos 20.000 genes aproximadamente, lo que le convierte en una de las especies con mayor número de genes en comparación con otros hongos (Martin y Nehls, 2009).

El análisis del genoma de *L. bicolor* ha revelado la existencia de una amplia variedad de enzimas hidrolíticas, además de proteasas, quitinasas y glucanasas extracelulares, que sirven para comprender mejor la capacidad que posee este hongo para actuar como un organismo saprófito facultativo. Por otro lado, existe un gran número de genes a los que todavía no se les ha asignado una función biológica. Entre ellos, cabe destacar algunos genes que aparentemente codifican lectinas y módulos de unión a carbohidratos (CBM). Principalmente, se han identificado lectinas, que por similitud de secuencia, quedan englobadas dentro de las superfamilias de las lectinas de tipo ricina, las galectinas y las lectinas de plantas leguminosas, como la ConA (<http://genome.jgi-psf.org/Lacbi2/Lacbi2.home.html>).

Las tres hipotéticas proteínas que podrían adquirir un plegamiento de tipo trébol  $\beta$  son las siguientes: 185716, 318163 y 587942. La proteína 185716 presenta similitud de secuencia con la lectina de *Xerocomus chrysenteron* y muestra especificidad por N-

acetilgalactósidos, tal y como se muestra anotado en la página web. La proteína 318163 coincide con la proteína B0D650\_LACBS, con un 52% de identidad de secuencia con el módulo lectina N-terminal de LSLa. Por último, de la proteína 587942 no hay información relevante acerca de posibles homólogos.

### 5.2.3. La lectina SSL de *Sclerotinia sclerotiorum*

*Sclerotinia sclerotiorum* es un hongo ascomiceto (orden Helotiales) patógeno de plantas con alimentación necrotrófica, que afecta a gran cantidad de plantas independientemente de su etapa de crecimiento. Supone un problema agrícola ya que muchos cultivos, como el de la patata, se ven afectados por el moho blanco, nombre con el que se conoce a la enfermedad de la que *S. sclerotiorum* es responsable. El hongo se caracteriza por la producción de unas estructuras oscuras, clave para su reproducción, denominadas esclerocios y compuestas por una masa de micelio endurecido, que en primavera fructifican y generan ascosporas (**Figura 1.20**).



**Figura 1.20.** Fotografía del hongo ascomiceto *Sclerotinia sclerotiorum* infectando una planta.

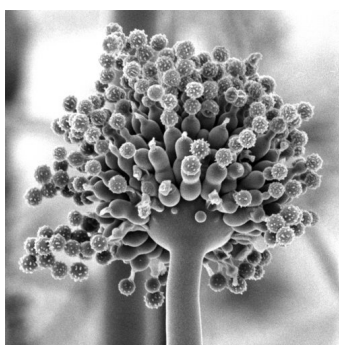
En el año 2011 se publicó el análisis del genoma de *S. sclerotiorum* (Amselem *et al.*, 2011). Su estudio intenta desvelar cómo se desencadena la patogenicidad en este hongo, sobre todo en lo referente a los mecanismos de infección. Junto con la caracterización de genes involucrados en la producción de peptidasas de acción extracelular y proteínas relacionadas con el metabolismo secundario, se han identificado algunos genes que codifican hipotéticas lectinas.

Resultados previos mostraron que al pasar extractos celulares solubles procedentes del hongo a través de una columna con D-galactosa inmovilizada, se obtiene, mediante elución con 0,1 M galactosa, un pico correspondiente a la aglutinina SSA (Kellens *et al.*, 1992; Sulzenbacher *et al.*, 2010). La lectina eluye pura y permanece como un dímero en solución. Al comparar el contenido en lectinas en diferentes estadios, se observa que éste aumenta desde un 0,2% (*p/p*) en el micelio a un 20% (*p/p*) en esclerocios cultivados durante dos

semanas (Kellens *et al.*, 1992). Aunque aparentemente, no se purifican más lectinas con esta metodología, los análisis genómicos indican que existen fragmentos génicos que codifican distintas lectinas, por lo que éstas podrían localizarse en otras partes del hongo o no ser funcionales en las condiciones experimentales en las que se trabajó en los casos anteriores. La hipotética lectina A7EMH9\_SCLS1, con un 34% de identidad con el módulo lectina de LSLa, solo muestra un 14% de identidad con la secuencia de la lectina SSA, aunque ambas son parecidas en tamaño y presentan una masa molecular de 17 kDa.

#### 5.2.4. La lectina ANL de *Aspergillus niger*

El hongo filamentoso *Aspergillus niger* (**Figura 1.21**) es uno de los más utilizados en las fermentaciones industriales para la producción de enzimas y ácidos orgánicos, particularmente ácido cítrico (Pandey *et al.*, 1999). Este hongo ascomiceto (orden Eurotiales) actúa en la naturaleza como un hongo parásito, produciendo el moho negro que aparece en frutas y vegetales. También puede ser un agente infeccioso para el hombre si es inhalado en grandes cantidades, produciendo alteraciones pulmonares conocidas como aspergilosis.



**Figura 1.21.** Micrografía del hongo ascomiceto *Aspergillus niger*. Este hongo es ampliamente utilizado en la industria para la producción, entre otros compuestos, de ácido cítrico.

El genoma de la cepa CBS 513.88 de *A. niger* (cepa usada en la industria para la producción de enzimas) fue secuenciado en 2007 (Pel *et al.*, 2007). En él se han identificado más de 14.000 marcos de lectura abiertos (ORF) dentro de las 33,9 Mpb. Muchos codifican gran cantidad de enzimas hidrolíticas, algunas de las cuales son secretadas al medio extracelular para ayudar al hongo a nutrirse degradando diferentes biopolímeros.

Aunque el análisis de la secuencia todavía no ha desvelado cuántas lectinas integran el genoma de *A. niger*, algunos ensayos bioquímicos ya han dado los primeros pasos. Hasta el momento se sabe que *A. niger*, al igual que otras especies del género *Aspergillus*, dispone solo de lectinas intracelulares con actividad hemaglutinante, que se inhiben en presencia de azúcares, como la N-acetilgalactosa, y glicoproteínas, como la mucina (Singh *et al.*, 2008). Por esta razón, la proteína A2R2Y9\_ASPNC probablemente sea una lectina intracelular del hongo.

## 6. Aplicaciones biotecnológicas de lectinas

A medida que se avanza en el conocimiento de las lectinas, es más fácil desarrollar aplicaciones biotecnológicas que exploten su actividad biológica en beneficio del hombre (Erjavec *et al.*, 2012). Uno de los primeros usos analíticos que se hicieron de las lectinas fue el desarrollo de una aplicación que ayudara a la tipificación de los antígenos que conforman los grupos sanguíneos del sistema ABO (Watkins, 1999). De esta manera, se observó, por ejemplo, cómo la lectina de *Anguilla anguilla* jugó un papel importante en demostrar que una característica exclusiva del grupo sanguíneo O es la presencia de residuos de fucosa (Watkins, 2001).

En la actualidad, las lectinas se usan como herramientas para el diagnóstico clínico, por ejemplo, en el análisis de patrones de glicosilación alterados en tumores malignos, permitiendo hacer un seguimiento de la progresión del tumor y de la metástasis. Los patrones de glicosilación, así como cualquier otro azúcar en general, se pueden analizar mediante micromatrices de lectinas o *lectin microarrays*<sup>27</sup>, basados en el uso de una placa en la que se disponen de forma ordenada lectinas inmovilizadas que exponen sus sitios de unión a azúcares (Petukhov *et al.*, 1999; Hirabayashi *et al.*, 2013). Una reacción de revelado indicará qué lectinas han unido azúcares y, por tanto, se podrá conocer la naturaleza del carbohidrato, valorando la especificidad de las lectinas en los pocillos que han dado positivo. Por otro lado, las lectinas también se usan ampliamente en la purificación de glicoconjugados. De esta manera, a partir de cualquier extracto celular, podemos separar la fracción que presenta un determinado azúcar incorporado. Como estas técnicas permiten distinguir entre una célula sana y otra enferma, las lectinas se usan activamente en el diseño racional de terapias y fármacos (Rahaie y Kazemi, 2010).

Existen multitud de trabajos en los que se ha demostrado el efecto antitumoral de algunas lectinas. Parece ser que éstas tienen la capacidad de inducir apoptosis en determinadas líneas celulares cancerígenas, activándose diferentes cascadas de señalización, como las que involucran a las caspasas. Así, se ha visto que la hemaglutinina de *Pseudomonas aeruginosa* actúa sobre diferentes cánceres de mama (Liu *et al.*, 2009) o que la lectina de *Polygonatum cyrtoneura* reconoce células A375 de melanoma humano (Liu *et al.*, 2009). Además de la actividad antitumoral, se ha descrito actividad antiviral. Se han encontrado algunas lectinas que se unen a la glicoproteína de superficie gp120 del virus VIH, inhibiendo la infección de células CD4 (Hansen *et al.*, 1989). Gracias a este tipo de resultados, el tratamiento del SIDA con lectinas está siendo profundamente investigado (Lam y Ng, 2011).



La agricultura también se beneficia del uso de las lectinas. Los nemátodos parásitos causan grandes pérdidas económicas en las cosechas y en el ganado. El descubrimiento de lectinas procedentes de hongos superiores con actividades nematotóxicas (Zhao *et al.*, 2009; Bleuler-Martínez *et al.*, 2011), que reconocen azúcares muy específicos, puede servir de arma para el desarrollo de vacunas efectivas para proteger al ganado, así como para el desarrollo de pesticidas que aseguren la perduración de las cosechas. Aunque representan una proporción mas pequeña, también existe un grupo que exhibe actividad antifúngica. Por ejemplo, la expresión de la lectina de *Gastrodia elata* en las células vasculares de la raíz y el tallo es fuertemente inducida por la presencia del hongo *Trichoderma viride*, lo que indica que la lectina es una proteína que forma parte del sistema de defensa de la planta (Sa *et al.*, 2009). Por tanto, la producción de plantas transgénicas que contuvieran estas lectinas podría protegerlas contra ataques de hongos parásitos.

A diferencia del desarrollo que ha experimentado la aplicación de las lectinas en el campo de la Biomedicina, hay todavía áreas hoy en día en las cuáles el uso de lectinas no representa una alternativa. Por esta razón, a continuación se va a introducir brevemente una aplicación en la que las lectinas fúngicas con plegamiento tipo trébol  $\beta$  podrían suponer, desde nuestro punto de vista, una herramienta tecnológica de interés: la producción de proteínas de fusión<sup>31</sup>.

## **6.1. Estandarización de la producción de proteínas a pequeña y gran escala**

Actualmente, se está dedicando un gran esfuerzo al desarrollo de protocolos eficaces de producción y purificación de proteínas recombinantes, tanto en investigación básica, a una escala de laboratorio medio, como en la búsqueda, mediante procesos de alto rendimiento (*high throughput*), de proteínas con potenciales aplicaciones terapéuticas, diagnósticas o enzimas industriales (Korf *et al.*, 2005; Arnau *et al.*, 2006). A ello contribuye, sin duda, la disponibilidad de una vasta cantidad de información genética, consecuencia del éxito de la *era genómica* (Abby y Daubin, 2007) y de técnicas eficaces de clonación y expresión de proteínas recombinantes (Hartley, 2006). La gran cantidad de información genética, especialmente proveniente de organismos procariotas, en parte se ha debido a la homogeneidad química y estructural de los ácidos nucleicos, lo cual ha permitido desarrollar protocolos de secuenciación de alto rendimiento de carácter universal. Por el contrario, es obvio que el



heterogéneo y complejo comportamiento de las proteínas en solución imposibilita *de facto* el establecimiento de protocolos genéricos de producción/purificación (Waugh, 2005).

La aproximación experimental que permite salvar la barrera impuesta por la heterogeneidad de las proteínas a la hora de desarrollar protocolos genéricos de producción-purificación es el uso de etiquetas de fusión<sup>12</sup> (*fusion tags*): péptidos o proteínas que incorporados a la proteína de interés formando una proteína de fusión, idealmente facilitarían su purificación y aumentarían el rendimiento de su producción sin alterar su estructura tridimensional ni su actividad (Terpe, 2003). A falta de una comparación sistemática (a una escala proteómica) de la eficacia de las principales etiquetas de fusión conocidas (**Tabla 1.4**), se puede afirmar que actualmente no hay una etiqueta “universal” (Esposito y Chatterjee, 2006): cada una presenta ventajas e inconvenientes. Es deseable, por lo tanto, la disposición de una amplia batería de ellas.

*Principales etiquetas de afinidad*

Etiqueta	Proteína o péptido	Matriz	Referencia
His <sub>6</sub>	Etiqueta de seis histidinas	Metales inmovilizados	(Gaberc-Porekar y Menart, 2001)
GST	Glutación-S-transferasa	Glutación	(Smith y Johnson, 1988)
MBP	Proteína de unión a maltosa	Amilosa	(di Guan <i>et al.</i> , 1988)
FLAG	Péptido FLAG	Anticuerpo anti-FLAG	(Einhauer y Jungbauer, 2001)
BAP	Péptido aceptor de biotina	Avidina	(Schatz, 1993)
Strep II	Péptido de unión a estreptavidina	Estreptavidina	(Voss y Skerra, 1997)
CBP	Péptido de unión a calmodulina	Calmodulina	(Vaillancourt <i>et al.</i> , 2000)

*Principales etiquetas solubilizantes*

Etiqueta	Proteína o péptido	Origen	Referencia
MBP	Proteína de unión a maltosa	<i>Escherichia coli</i>	(Nallamsetty y Waugh, 2006)
GST	Glutación-S-transferasa	<i>Schistosoma japonicum</i>	(Nygren <i>et al.</i> , 1994)
Trx	Tiorredoxina	<i>Escherichia coli</i>	(LaVallie <i>et al.</i> , 1993)
NusA	<i>N utilization substance A</i>	<i>Escherichia coli</i>	(Davis <i>et al.</i> , 1999)
SUMO	Pequeño modificador de ubiquitina	<i>Homo sapiens</i>	(Marblestone <i>et al.</i> , 2006)
SET	Etiqueta potenciadora de la solubilidad	Sintético	(Zhang <i>et al.</i> , 2004)

**Tabla 1.4.** Principales etiquetas de fusión utilizadas en la expresión y purificación de proteínas recombinantes.

Hoy en día el interés se centra principalmente en MBP, GST, NusA y Trx. Sin embargo, teniendo en cuenta que estudios recientes demuestran una baja eficacia de GST (Dyson *et al.*, 2004) y Trx (Marblestone *et al.*, 2006) como agentes solubilizantes (*solubility enhan-*

cers) en *E. coli* y que NusA, aun con una eficacia similar a MBP como agente solubilizante, no permite desarrollar un protocolo de purificación sencillo (Davis *et al.*, 1999; Dummer *et al.*, 2005), MBP se ha convertido en una de las etiquetas de fusión sobre la que más interés se tiene y una de las más ampliamente estudiadas (Dyson *et al.*, 2004; Braud *et al.*, 2005).

Una etiqueta de fusión ideal aumentaría el rendimiento en la producción de proteínas. Este efecto está asociado normalmente con etiquetas proteicas (no péptidos pequeños) situadas en el extremo N-terminal de la fusión, en donde la construcción proporcionaría un entorno eficaz para iniciar el proceso de traducción, evitando que el ARNm adopte estructuras secundarias que pudiesen interferir en la unión a los ribosomas (Hartz *et al.*, 1991). Asimismo, dichas etiquetas también deben permitir desarrollar un protocolo de purificación genérico, eficaz y potencialmente escalable a procesos de alto rendimiento con costes mínimos. En este sentido, la matriz cromatográfica empleada cobra especial relevancia, distinguiéndose dos grandes grupos: matrices con proteínas inmovilizadas, que suelen ser caras, con baja capacidad de unión y vidas medias cortas, y matrices con ligandos de pequeño tamaño, más baratas y estables. Por otro lado, la disociación de la proteína inmovilizada en la matriz debe producirse en condiciones no agresivas para el estado nativo de la proteína de interés.

Claramente, cualquier estrategia basada en el uso de etiquetas de fusión debe proporcionar la posibilidad de eliminarla de la construcción de un modo eficaz, aspecto este especialmente relevante cuando se trata de proteínas con potenciales aplicaciones en terapia humana (Arnau *et al.*, 2006). La disponibilidad hoy día de endoproteasas con una especificidad exquisita, como la proteasa del virus del grabado del tabaco (TEV del inglés *Tobacco Etch Virus*), la enteroquinasa o la enzima PreScission (GE Healthcare Life Sciences), ha mitigado notablemente este problema. De este modo, la construcción de la proteína de fusión incluye una secuencia de reconocimiento por una de estas proteasas situada entre la etiqueta de fusión y la proteína de interés. Puesto que los determinantes de la especificidad de estas endoproteasas se localizan en el lado carbonilo del enlace peptídico diana, tras la hidrólisis, estos permanecerán en las etiquetas eliminadas en el caso de estar situadas hacia el extremo N-terminal de la proteína de fusión.

De acuerdo con las expectativas de aumento de inversiones en el mercado farmacéutico en fármacos de naturaleza proteica, se espera la aparición de nuevas proteínas de interés terapéutico. De ahí que la búsqueda de nuevas etiquetas de afinidad sea actualmente un campo de intensa investigación (Arnau *et al.*, 2006).

## 6.2. Lectinas empleadas como etiquetas de fusión

Actualmente, son pocas las proteínas con capacidad para unir carbohidratos empleadas como etiquetas de fusión. En general, podemos distinguir tres grupos: dominios de unión a celulosa (Tomme *et al.*, 1998), dominios de unión a quitina y lectinas. Por un lado, los dominios de unión a celulosa o CBD (del inglés *Cellulose-Binding Domains*) constituyen un grupo muy heterogéneo. Existen más de 13 familias diferentes cuyo tamaño varía desde los 4 hasta los 20 kDa. La unión a celulosa puede ser tanto reversible como irreversible, por lo que en este último lugar, su aplicación se dirige exclusivamente a la inmovilización de enzimas<sup>19</sup> (Xu *et al.*, 2002). Cuando la unión es reversible, los CBD se utilizan como etiquetas de afinidad, siendo los miembros de las familias I, II y III los más usados. Los CBD de la familia I presentan una unión más fuerte, por lo que las condiciones de elución son drásticas, requiriéndose soluciones desnaturizantes con urea o cloruro de guanidinio, lo que implica que la purificación incluya obligatoriamente una etapa de replegamiento de la proteína de fusión. Sin embargo, los CBD de las otras dos familias se unen menos intensamente y la elución se puede llevar a cabo con etilenglicol, un agente menos agresivo que puede ser eliminado fácilmente mediante una diálisis (McCormick y Berg, 1997). El soporte al que se unen estas proteínas, la celulosa, presenta varias ventajas, destacándose que es inerte, no experimenta uniones inespecíficas y su uso ha sido aprobado por muchas empresas farmacéuticas para la producción de fármacos dirigidos a seres humanos (Terpe, 2003).

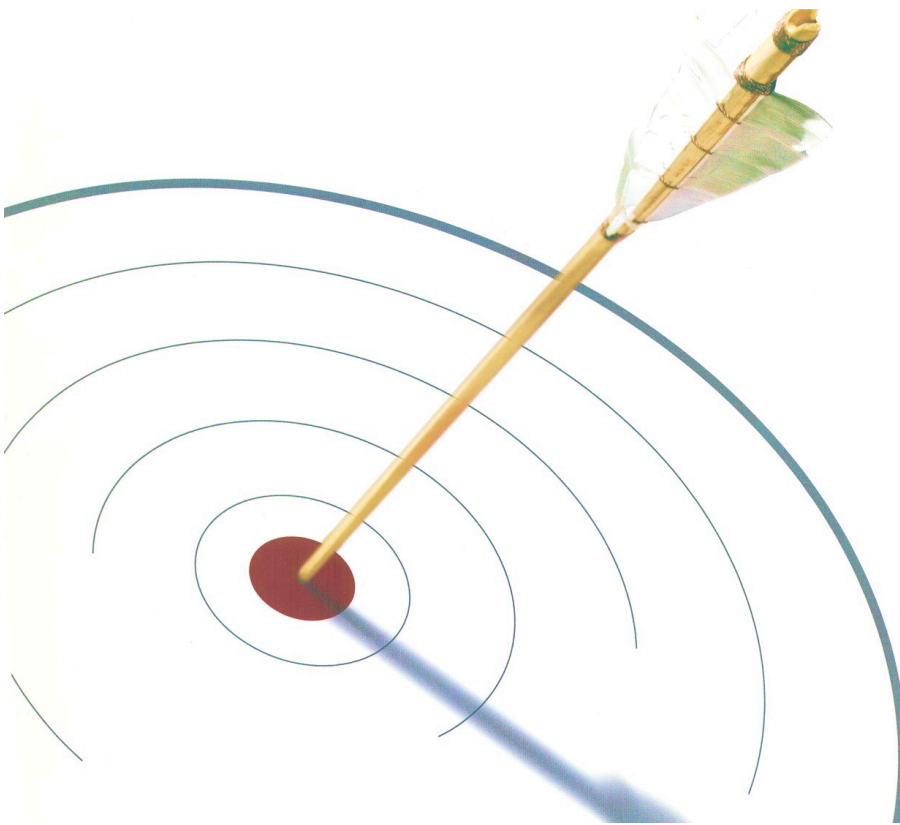
En segundo lugar, destacan los dominios de unión a quitina, conocidos en la literatura de nuevo como CBD (del inglés *Chitin-Binding Domains*). Uno de los dominios de unión a quitina más ampliamente utilizado es el procedente de la bacteria *Bacillus circulans* de tan solo 51 aminoácidos (Watanabe *et al.*, 1994). Este tipo de etiquetas de fusión suele utilizarse en combinación con inteínas, que tienen la capacidad de escindirse por sí mismas de la proteína precursora (*self-splicing*). La reacción de autoprosesado es inducida mediante el uso de agentes reductores, como 1,4-ditiotreitol (DTT) o  $\beta$ -mercaptoetanol. El dominio de unión a quitina puede eluirse de la columna en condiciones no fisiológicas, por ejemplo, con un 1% de SDS o con 6 M cloruro de guanidinio (Terpe, 2003).

Por último, las lectinas también constituyen un grupo de etiquetas de fusión, aunque sólo se han desarrollado unos pocos sistemas basados en ellas. Una de las lectinas más importantes es la proteína de unión a maltosa o MBP (del inglés *Maltose-Binding Protein*), que como se ha visto en el apartado anterior, es una de las etiquetas más ampliamente utilizadas para producir y purificar proteínas recombinantes en bacterias. MBP es una proteína

de 43 kDa, codificada por el gen *malE* de *E. coli* K12 (Duplay *et al.*, 1988), que se expresa con un rendimiento muy elevado. Las proteínas de fusión que incorporan MBP como etiqueta pueden acomodarla tanto en el extremo N-terminal como en el C-terminal (Dyson *et al.*, 2004). De esta forma, estas proteínas de fusión pueden purificarse en un único paso cromatográfico mediante cromatografía de afinidad de unión a resinas de amilosa. La elución se lleva a cabo de forma sencilla, con una solución conteniendo maltosa en un tampón con un pH 7,0-8,5 y hasta una concentración 1 M de sal. MBP no sólo actúa de etiqueta de afinidad, sino que incrementa la solubilidad de la proteína diana a la que está fusionada (Kataeva *et al.*, 2005). Esto hace que MBP se utilice como una alternativa para conseguir solubilizar proteínas que tienden a formar agregados o cuerpos de inclusión ya que no interfiere en el plegamiento de la proteína diana (la proteína de fusión estaría formada por dos módulos independientes de plegamiento). El sistema de purificación basado en MBP presenta ciertos inconvenientes relacionados con la resina. Por ejemplo, la amilosa no puede ser empleada en presencia de agentes reductores y/o desnaturalizantes. Además, los cultivos deben ser suplementados con 0,2% (*p/v*) glucosa para evitar que la amilosa sea degradada por la actividad amilasa de *E. coli* (Malhotra, 2009).

Junto con MBP, hay otras lectinas que también se han aplicado como etiquetas de fusión para la expresión-purificación de proteínas recombinantes. Un ejemplo de ello es la proteína LecB de *Pseudomonas aeruginosa*, una lectina tetramérica de 11,73 kDa por monómero y específica de L-fucosa y D-manosa, cuyo rendimiento de producción en *E. coli* es bastante elevado (70 mg/L) (Gilboa-Garber *et al.*, 2000). La purificación de la proteína de fusión YFP-LecB (con la lectina en el extremo C-terminal y un sitio de corte para enteroquinasa), llevada a cabo mediante un único paso cromatográfico, empleando una columna de manosa-agarosa y eluyendo con 5-20 mM maltosa, demostró que la lectina actúa como una buena etiqueta de afinidad (Tielker *et al.*, 2006). En la bibliografía, se encuentran otros ejemplos, como el de la congerina II (una galectina de congrio), que se utilizó para producir recombinantemente una fosfolipasa de serpiente (Seto *et al.*, 2008), o el desarrollo de una nueva estrategia de purificación de proteínas recombinantes en plantas basada en el uso de la aglutinina de la soja (SBA) como etiqueta de afinidad (Tremblay *et al.*, 2011).

# CAPÍTULO 2



## **OBJETIVOS**



## II

## OBJETIVOS

## 1. Hipótesis de partida y objetivos

Teniendo en cuenta la información descrita anteriormente en la *Introducción*, se puede concluir que la investigación de lectinas fúngicas con estructura tipo trébol  $\beta$  podría ayudarnos a interpretar mejor cómo interactúan los hongos en su entorno natural, así como a desarrollar novedosas aplicaciones biotecnológicas de interés. La solicitud de una patente española a la Oficina Española de Patentes y Marcas de la optimización de un sistema de expresión de proteínas de fusión con el módulo lectina N-terminal de LSLa como etiqueta de afinidad es un claro ejemplo que justifica el potencial intrínseco de estos dominios lectina con estructura trébol  $\beta$ . La patente se encuentra, de hecho, en fase de comercialización desde el 2 de abril de 2008: “Proteínas de fusión con un dominio lectina de tipo trébol beta, procedimientos de obtención y sus aplicaciones” (número: 200800909).

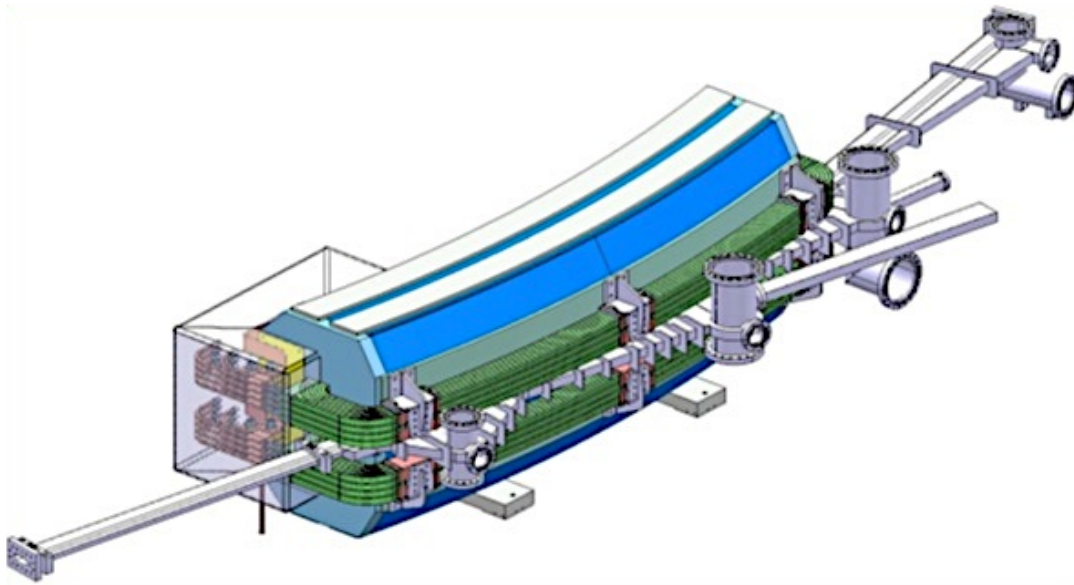
Esta tesis doctoral, por tanto, parte de la siguiente hipótesis: los análisis estructurales y funcionales de dominios lectina con estructura tipo trébol  $\beta$  proporcionarán información útil para la identificación de su función biológica y una base racional para el desarrollo de potenciales aplicaciones biotecnológicas. En base a esta hipótesis, se han establecido los siguientes objetivos para la presente tesis doctoral:

1. Clonación, expresión, purificación y caracterización estructural de las proteínas homólogas al dominio lectina N-terminal de LSLa identificadas en los organismos *Laccaria bicolor*, *Sclerotinia sclerotiorum* y *Aspergillus niger*.
2. Determinación de la especificidad de las lectinas e identificación de su posible función biológica.
3. Obtención de variantes mejoradas del módulo lectina N-terminal de LSLa y estudio detallado de las interacciones lectina-azúcar.
4. Desarrollo de novedosas aplicaciones biotecnológicas mediante la aplicación de los módulos lectina con estructura tipo trébol  $\beta$ .

La consecución de los objetivos planteados implica la aplicación de tres aproximaciones experimentales distintas: (1) técnicas de biología molecular, (2) técnicas bioquímicas y biofísicas, y (3) técnicas de cristalografía de proteínas. Cabe destacar que el logro del primer objetivo es determinante para el posible uso de las proteínas en diferentes aplicaciones biotecnológicas. De esta manera, la ausencia de interacción con matrices derivadas de agarosa invalidaría a estas proteínas como etiquetas de afinidad y no servirían para el desarrollo de aplicaciones biotecnológicas. Esta prueba sirve, por tanto, como criterio de suficiencia.



# CAPÍTULO 3



## **MATERIALES Y MÉTODOS**



## III

## MATERIALES Y MÉTODOS

## 1. Visión global

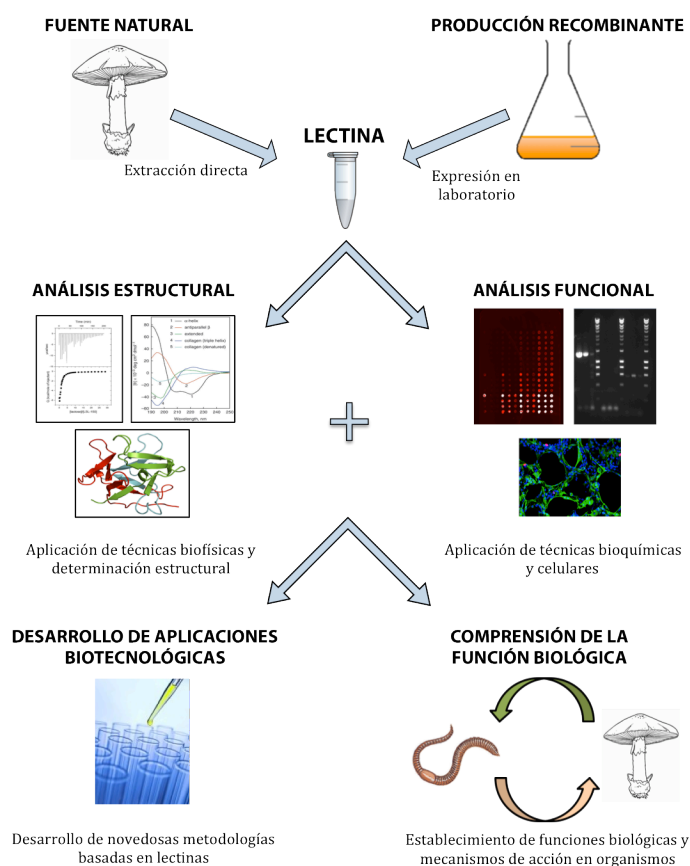
En términos generales, un análisis exhaustivo enfocado en el estudio de proteínas y de lectinas, en particular, debería plantearse desde una perspectiva multidisciplinar, esto es, combinando diferentes estrategias experimentales (**Figura 3.1**). Esta aproximación metodológica es hoy en día la manera más eficaz para comprender el mecanismo de acción de estas macromoléculas y, en este sentido, se puede afirmar que el papel que juega el conocimiento de sus estructuras tridimensionales es fundamental, ya que constituye el marco de referencia para explicar otras propiedades, así como las interacciones que establecen con otras moléculas, azúcares en el caso de las lectinas. Esto, a su vez, nos permitirá comprender su especificidad (Gabius *et al.*, 2011), clave para llevar a cabo su función biológica y de gran utilidad en el desarrollo de aplicaciones biotecnológicas novedosas.

A continuación, se van a detallar todos los procedimientos experimentales, así como a describir las diferentes técnicas utilizadas en el desarrollo de este trabajo, lo cual ha comprendido la producción de las lectinas fúngicas de modo recombinante, su caracterización estructural y analítica así como las aplicaciones biotecnológicas desarrolladas.

## 2. Expresión y purificación de lectinas fúngicas recombinantes

A pesar de que las lectinas de plantas y animales han sido extensamente estudiadas (Rini y Lobsanov, 1999; Rudiger y Gabius, 2001), el estudio de las lectinas fúngicas ha atraído la atención recientemente y, en concreto, en la última década, éste se ha visto fuertemente incrementado (Khan y Khan, 2011; Varrot *et al.*, 2013). Por lo general, estas lectinas se obtienen directamente de la fuente natural: micelio (Candy *et al.*, 2003), cuerpos fructíferos (Tateno y Goldstein, 2003), esporas (Tronchin *et al.*, 2002) y/o medio extracelular (Bhowal *et al.*, 2005). Sin embargo, gracias al desarrollo de las técnicas modernas de Biología Molecular, la obtención de lectinas recombinantes, con alto rendimiento, ha facilitado su estudio.

Por regla general, la mayoría de lectinas se purifican mediante métodos cromatográficos. Aunque en la bibliografía podemos encontrar lectinas fúngicas que se purifican usando columnas de intercambio iónico (Zhao *et al.*, 2009) o exclusión molecular (Tronchin *et al.*, 2002), la mayor parte de ellas se han purificado mediante cromatografía de afinidad, puesto que es una metodología mucho más selectiva. Por esta razón, éste último ha sido el método elegido para la purificación de las proteínas caracterizadas en este trabajo.



**Figura 3.1.** Diagrama que muestra las diferentes aproximaciones experimentales en el estudio exhaustivo de una lectina.

## 2.1. Obtención de la lectina LSLa de *Laetiporus sulphureus* y de su correspondiente módulo C-terminal formador de poros

Se produjeron de forma recombinante tanto la proteína completa LSLa como su correspondiente módulo C-terminal formador de poros (PFM). Ambos fragmentos génicos se amplificaron mediante PCR (reacción en cadena de la polimerasa) a partir del plásmido pET-LSLa (Tateno y Goldstein, 2003) (cedido por H. Tateno e I. J. Goldstein) utilizando los pares de oligonucleótidos 1/4 y 3/4 (**Tabla 3.1**), respectivamente. La clonación en el vector de

expresión pET-28a(+) se llevó a cabo con las enzimas de restricción *NcoI* y *HindIII* (New England Biolabs), obteniéndose como resultado los correspondientes plásmidos pKLSLa y pKPfM, que contienen el promotor del fago T7 y un gen de resistencia a kanamicina.

Los plásmidos fueron utilizados para transformar células competentes de *Escherichia coli* DH5 $\alpha$ , así como *E. coli* BL21(DE3) (Novagen). Mientras que las primeras células se utilizaron para producir cantidades adicionales de plásmido, las segundas, se emplearon para la producción de proteína recombinante. Las células transformadas mediante choque térmico a 42 °C fueron sembradas en placas de LB-agar (Sambrook y Russell, 2001) con kanamicina (50  $\mu$ g/mL) (LB-Km). Después de una incubación a 37 °C durante toda la noche, se obtuvieron colonias aisladas que contenían los plásmidos de interés.

Para la producción de la proteína LSLa y del módulo PFM, empleando células de *E. coli* BL21(DE3) previamente transformadas con los correspondientes plásmidos, se prepararon preinóculos de 10 mL de cultivo en medio LB-Km. Estos se dejaron crecer toda la noche a 37 °C con agitación. Una vez crecidos, cada cultivo se transfirió a 1 L de medio LB-Km en matraces *Erlenmeyer* de 2 L de capacidad. Cuando la densidad óptica del medio a una longitud de onda de 600 nm (DO<sub>600</sub>) alcanzó valores en torno a 0,4-0,6 (fase exponencial de crecimiento del cultivo), se indujo la expresión de proteína recombinante, añadiendo isopropiltiogalactósido (IPTG; concentración final: 0,4 mM), un potente inductor del operón lactosa. Después de la inducción, las células se incubaron a 16 °C durante aproximadamente unas 20 horas adicionales. Finalmente, los cultivos celulares se centrifugaron a 4.500 rpm (IEC CL40 Centrifuge, Thermo Scientific) durante 20 minutos para preparar los precipitados celulares que se conservaron a -80 °C en espera de ser procesados.

Para la purificación de LSLa y PFM, en primer lugar, se resuspendieron los precipitados celulares en tampón de lisis TSA (20 mM Tris-HCl pH 8,0, 100 mM NaCl y 0,04% (p/v) NaN<sub>3</sub>) en presencia de un cóctel de inhibidores de proteasas (Roche), haciendo uso de 15-20 mL de tampón por litro de cultivo. A continuación, las células se lisaron sometiendo a alta presión con una *French Press* (Thermo Scientific). Los lisados obtenidos se centrifugaron a 40.000  $\times$  g durante 30 minutos a 4 °C en un rotor SS34 con una centrífuga SORVALL Evolution RC (Kendro) para separar las fracciones soluble e insoluble del extracto celular. Estas fracciones fueron evaluadas mediante SDS-PAGE para comprobar qué fracción estaba más enriquecida en proteína recombinante y, así, estimar cualitativamente la solubilidad de las proteínas LSLa y PFM.

Por otro lado, el extracto celular de LSLa se cargó directamente a un flujo de 1 mL/min en una columna preparada manualmente (Glass Econo-Column® de 2,5 × 10 cm, Bio-Rad) de Sepharose® 4B (Sigma-Aldrich) preequilibrada en tampón TSA. La cromatografía se llevó a cabo con un sistema cromatográfico BioLogic LP (Bio-Rad). Después de un lavado exhaustivo con tampón TSA, LSLa se eluyó con tampón de elución TSAL (20 mM Tris-HCl pH 8,0, 100 mM NaCl, 0,04% (p/v) NaN<sub>3</sub> y 0,2 M lactosa). La pureza de LSLa se analizó mediante SDS-PAGE.

## **2.2. Obtención del módulo lectina N-terminal de la proteína LSLa de *Laetiporus sulphureus***

La producción del módulo lectina N-terminal (LSL<sub>150</sub>) de la proteína LSLa de *Laetiporus sulphureus* comenzó con la clonación del fragmento génico correspondiente. El segmento codificante para la secuencia de los primeros 150 aminoácidos de LSLa fue amplificado mediante PCR con el par de oligonucleótidos 1/2 (**Tabla 3.1**) y usando como molde el vector pET-LSLa (Tateno y Goldstein, 2003). El fragmento amplificado se clonó en el vector pET-28a(+) (Novagen) mediante el uso de las enzimas de restricción *NcoI* y *EcoRI* (New England Biolabs). Como resultado, se obtuvo el vector pKLSL<sub>150</sub> (Angulo *et al.*, 2011).

La expresión de las proteínas se llevó a cabo en células de *E. coli* BL21(DE3) (Novagen), siguiendo el mismo protocolo aplicado en el caso anterior. Los precipitados celulares obtenidos después de la inducción se lisaron con ayuda de una *French Press* (Thermo Scientific) y posteriormente se centrifugaron a 40.000 × g durante 30 minutos a 4 °C. La fracción soluble se filtró con un filtro de celulosa de 0,45 µm de diámetro de poro, quedando la muestra del extracto celular preparada para la purificación.

LSL<sub>150</sub> se purificó mediante cromatografía de afinidad a Sepharose® 4B (Sigma-Aldrich), siguiendo el protocolo descrito en el apartado anterior 2.1 para LSLa. Al terminar el proceso de purificación, la proteína se concentró mediante ultrafiltración hasta una concentración final de ~10 mg/mL y se almacenó a -80 °C. La concentración de lectina se halló mediante un método espectroscópico usando un espectrofotómetro ND-1000 (NanoDrop) y empleándose un coeficiente de absorción  $E^{0,1\%}$  (1 cm, 280 nm) de 1,496, el cual fue determinado experimentalmente.

Número (Sentido)	Gen que se amplifica	Vector de clonación	Secuencia 5'→3'	Vector
1 (D)	fragmento <i>lsl<sub>150</sub></i> y <i>lsla</i>	pET-28a(+)	gatg <b>ccatgg</b> caATGACCGACATC	pKLSL <sub>150</sub> y pKLSLa
2 (R)	fragmento <i>lsl<sub>150</sub></i>	pET-28a(+)	<b>cgaattc</b> TCAATCCTCGAAGA	pKLSL <sub>150</sub>
3 (D)	fragmento <i>pfm</i>	pET-28a(+)	catg <b>ccatgg</b> GCATCGATAAGATAGAGTACGAC	pKPFM
4 (R)	fragmento <i>pfm</i> y <i>lsla</i>	pET-28a(+)	ccc <b>aagctt</b> TCAGGCCGTTACCGAAGTCAGTGTATG	pKPFM y pKLSLa
5 (D)	fragmento <i>lbl<sub>152</sub></i> y <i>lbl</i>	pET-28a(+)	gatg <b>ccatgg</b> CAAACGAGTACAATCCCCCTCTTG	pKLBL y pKLBL <sub>152</sub>
6 (R)	<i>lbl</i>	pET-28a(+)	<b>ggaattc</b> TTAAAGACACCAATAACAGC	pKLBL
7 (R)	fragmento <i>lbl<sub>152</sub></i>	pET-28a(+)	<b>ggaattc</b> TCAGTCGAATAAGATAGTGAAATATTG	pKLBL <sub>152</sub>
8 (R)	<i>lblexp</i>	pET-28a(+)	<b>ggaattc</b> TCAATGAGGAACCCGG	pKLBLexp
9 (D)	<i>lblexp</i>	YEp-FLAG	acgcct <b>cgga</b> TGTCAAACGAGTACAATCCCC	YEp-LBL
10 (R)	<i>lblexp</i>	YEp-FLAG	cc <b>ggaattc</b> TCAATGAGGAACCCGGAATGCAG	YEp-LBL
11 (D)	<i>lblexp</i>	YEp-FLAG	<i>gactacaaggatgacgatgacaag</i> ATGTCAAACGAGTACAATCCCCCTC	YEpFLAG-LBL
12 (R)	<i>lblexp</i>	YEp-FLAG	<i>ggttttgggacgctcgacggatc</i> TCAATGAGGAACCCGGAATGCAGCG	YEpFLAG-LBL
13 (D)	<i>ssl</i>	pKHLSLt	<i>tgaaaacctgtatttccagggc</i> ATGGTTGTTGAAGTCCGAGTCCGG	pKHLSLt-SSL
14 (R)	<i>ssl</i>	pKHLSLt	<i>tcagcttcctttcgggctttgtta</i> TTAGGTAAAGGTTTGAATCTG	pKHLSLt-SSL
15 (D)	<i>anl</i>	pURI3-Cter	<i>taactttaagaaggagatatatacat</i> ATGCCGATCCCGACCGACGCG	pURI3-ANLHis
16 (R)	<i>anl</i>	pURI3-Cter	<i>gctattaatgatgatgatgatgatg</i> GTCATTTCAGGTAATACAGCACC	pURI3-ANLHis

**Tabla 3.1.** Secuencias de los oligonucleótidos utilizados para la construcción de los diferentes vectores de expresión de lectinas fúngicas. En *minúscula y negrita* aparecen los sitios de reconocimiento de las enzimas de restricción. En *minúscula y cursiva* se muestran los sitios que hibridan directamente con el plásmido para la clonación independiente de enzimas de restricción y ligasa. En *mayúscula* se indican las regiones que hibridan con el inserto. Las letras D y R indican el sentido de lectura del oligonucleótido: D significa sentido directo (5'→3') y R, sentido reverse (3'→5' de la hebra codificante).

### 2.3. Obtención de la lectina LBL de *Laccaria bicolor* y de su correspondiente módulo lectina N-terminal

Utilizando la secuencia de aminoácidos de LSL<sub>150</sub> se llevó a cabo una búsqueda de proteínas homólogas utilizando el programa BLAST (*Basic Local Alignment Search Tool*) del servidor ExpASY (Gasteiger *et al.*, 2003). El resultado que mostraba una mayor identidad de secuencia se correspondía con el extremo N-terminal de una proteína procedente del hongo *Laccaria bicolor* (código UniProtKB: B0D650\_LACBS) con función biológica desconocida.

Para la caracterización de la hipotética lectina, en primer lugar, se procedió a la clonación tanto de la secuencia que codificaba la proteína completa (gen *lbl* cuyo código es el LACBIDRAFT\_318163) como de la correspondiente al fragmento codificante de los 152 primeros aminoácidos (*lbl*<sub>152</sub>, versión corta N-terminal del gen *lbl*), que correspondería previsiblemente al módulo lectina N-terminal. El material biológico del que se partió para llevar a cabo las correspondientes clonaciones consistía en diferentes bibliotecas de ADN codificante (ADNc) que fueron suministradas por el grupo de investigación del Dr. Francis Martin del *Institut de la Recherche Agronomique* (Nancy, Francia). Se trabajó con tres bibliotecas diferentes, todas ellas de la cepa *L. bicolor* S238N. Una de las bibliotecas procedía de la ectomicorriza entre un espécimen de invernadero de tres meses y un álamo de la especie *Populus trichocarpa* (ECM-I), mientras que las otras dos se obtuvieron de los cuerpos fructíferos en diferentes estadios de crecimiento del hongo: FB-I (estadio temprano) y FB-II (estadio tardío). En concreto, para la obtención de las dos últimas bibliotecas, el hongo fue inoculado en un abeto de Douglas y crecido controladamente en invernadero (Villeneuve *et al.*, 1991). La extracción del ARN fue llevada a cabo utilizando el RNeasy Plant Mini Kit (Qiagen) para posteriormente ser amplificado y transcrito a ADNc haciendo uso del SMART PCR cDNA Synthesis Kit (Clontech).

#### 2.3.1. Producción de la lectina completa de *Laccaria bicolor*

La producción de la lectina completa, definida aquí como LBL (del inglés *Laccaria bicolor Lectin*), se llevó a cabo empleando dos sistemas de expresión distintos: bacterias y levaduras. En ambos casos, el gen *lbl* tuvo que ser amplificado para su posterior inserción en un vector de expresión. La información acerca de los oligonucleótidos utilizados en las correspondientes clonaciones queda recogida en la **Tabla 3.1**. Cabe resaltar que los resultados de secuenciación (apartado 3.2 de *Materiales y Métodos*) indicaron que el gen *lbl* presentaba una secuencia diferente a la mostrada en la base de datos pública del genoma del



hongo, por lo que se optó por clonar la secuencia determinada experimentalmente. Este nuevo gen, denominado *lblexp* (secuencia experimental del gen *lbl*), fue el utilizado en las distintas estrategias de expresión.

Se eligieron las células de *E. coli* BL21(DE3) (Novagen) como sistema de expresión. Además de seguir el protocolo general de expresión y purificación aplicado para LSL<sub>150</sub>, se llevó a cabo una coexpresión con la chaperona de *E. coli* GroEL/GroES (Thomas *et al.*, 1997). Para ello, las células de *E. coli* BL21(DE3) fueron cotransformadas con los plásmidos pKLBLe<sub>x</sub>p y pGro7 (Takara) y crecidas bajo una doble selección en medio LB con kanamicina (50 µg/mL) y cloranfenicol (20 µg/mL) (medio LB-Km-Cm). Los cultivos se prepararon con L-arabinosa (2 mg/mL, Sigma-Aldrich), ya que actúa como inductor de la sobreexpresión de chaperona. Cuando las células alcanzaron la fase exponencial, se añadió IPTG (0,4 mM) para inducir la producción de la lectina recombinante.

La otra alternativa de producción recombinante consistió en utilizar la levadura *Saccharomyces cerevisiae*, concretamente, células BJ3505, caracterizadas por presentar auxotrofia frente a triptófano y ser deficientes en proteasas. Estas células fueron transformadas usando el kit Frozen-EZ Yeast Transformation II (Zymo Research) con los plásmidos YEp-LBL y YEpFLAG-LBL. Ambos contienen el gen *lblexp*, aunque el vector YEpFLAG-LBL permite asimismo exportar la proteína al espacio extracelular gracias a la presencia del péptido señal *FLAG*. Aunque la selección de los clones se hizo en medio YNB-Trp<sup>-</sup>, las levaduras se crecieron en medio rico YPHSM (1% (p/v) glucosa, 1% (p/v) extracto de levadura, 8% (p/v) peptona y 3% (v/v) glicerol) a 30 °C y con agitación. La inducción de la producción de lectina, dependiente de la disminución de la concentración de glucosa en el medio, se mantuvo durante 72 horas a 30 °C, ya que fue el tiempo para el que se habían observado previamente mayores niveles de expresión. El protocolo seguido para la purificación de la lectina fue el mismo que el descrito arriba en la sección 2.1 de *Materiales y Métodos*, aunque con ciertas variaciones. La lisis celular en el caso de la producción intracelular se llevó a cabo con un agitador de bolas de vidrio (*bead beater*, GlenMills); mientras que en el caso de la producción extracelular, fue la fracción soluble obtenida tras recoger los precipitados celulares la que se cargó directamente en la columna de Sepharose® 4B para la cromatografía de afinidad.

### 2.3.2. Producción de LBL<sub>152</sub>: el módulo lectina N-terminal de la proteína LBL

La amplificación del fragmento *lbl<sub>152</sub>* se llevó a cabo utilizando los pares de oligonucleótidos 5/7 (Tabla 3.1). Siguiendo la misma estrategia descrita anteriormente para la construcción del vector pKLSL<sub>150</sub>, se clonó *lbl<sub>152</sub>* en el vector pET-28a(+) (Novagen), obteniéndose finalmente el correspondiente vector pKLBL<sub>152</sub>.

La expresión y la purificación del módulo lectina N-terminal (LBL<sub>152</sub>) se llevó a cabo siguiendo un protocolo equivalente al utilizado para LSL<sub>150</sub>. Se empleó un único paso de purificación en Sepharose® 4B (Sigma-Aldrich) como resina de afinidad. Finalmente, el cálculo de la concentración de proteína soluble se realizó espectroscópicamente usando un coeficiente de absorción  $E^{0,1\%}$  (1 cm, 280 nm) de 1,518, determinado experimentalmente.

## 2.4. Obtención de la lectina SSL de *Sclerotinia sclerotiorum*

La segunda lectina homóloga a LSL<sub>150</sub> que se estudió fue una proteína de 149 residuos del hongo *Sclerotinia sclerotiorum* (código UniProtKB: A7EMH9\_SCLS1). Esta proteína (SSL, del inglés *Sclerotinia sclerotiorum* *Lectin*) apareció en la búsqueda de homólogos con un porcentaje menor de similitud de secuencia que el observado en LBL. Al igual que esta última, se trata de una proteína que hasta el momento no ha sido caracterizada ni funcional ni estructuralmente.

La síntesis química del gen *ssl* que codifica SSL fue encargada a la empresa GenScript (Piscataway, Nueva Jersey). La secuencia de nucleótidos, además, fue optimizada con el fin de mejorar la expresión en *E. coli*. Tras comprobar su secuencia, el gen *ssl* fue clonado en los vectores de expresión pET-28a(+) (entre las dianas *NcoI* y *EcoRI*) y pKHLSLt (ver el apartado 7.1 de *Materiales y Métodos*). De esta manera, se produjeron dos variantes de la proteína SSL: una carente de etiqueta de afinidad y otra como una proteína de fusión con la etiqueta His-LSLt en su extremo N-terminal. En ambos casos, la purificación se llevó a cabo utilizando cromatografía de afinidad, tal y como se ha descrito anteriormente en la sección 2.1 de *Materiales y Métodos*.

## 2.5. Obtención de la lectina ANL de *Aspergillus niger*

Otra proteína homóloga a LSL<sub>150</sub>, identificada en la búsqueda llevada a cabo con el programa *BLAST*, procedía del hongo *Aspergillus niger* (código UniProtKB: A2R2Y9\_ASPNC). La

similitud de secuencia con respecto a LSL<sub>150</sub> es inferior a las de LBL y SSL, y como éstas, su función biológica es desconocida a día de hoy.

El gen *anl* que codifica la hipotética lectina de *A. niger* también fue sintetizado por el servicio de síntesis de la empresa GenScript. El gen se sintetizó con la secuencia optimizada para su expresión en *E. coli* y se insertó en el vector pUC57, que presenta resistencia a ampicilina. Este plásmido se usó como molde para amplificar el gen *anl* con el par de oligonucleótidos 15/16 (**Tabla 3.1**) y subclonarlo en el vector pURI3-Cter (Curiel *et al.*, 2011) con resistencia a ampicilina, que permite insertar una cola de seis histidinas en posición C-terminal. La clonación se llevó a cabo mediante un método independiente de enzimas de restricción y de ligasa (de las Rivas *et al.*, 2007).

La expresión de la lectina ANL (*Aspergillus niger* *Lectin*) se llevó a cabo en células de *E. coli* BL21(DE3). Éstas se crecieron en medio LB con ampicilina (100 µg/mL) (medio LB-Ap) a 37 °C hasta que alcanzaron la fase exponencial. Llegados a este punto, se indujo la sobreexpresión de proteína añadiendo IPTG (0,4 mM). Después de unas 20 horas de inducción a 22 °C, los precipitados celulares se recogieron mediante centrifugación a 4.500 rpm y 4 °C (IEC CL40 Centrifuge, Thermo Scientific), y se almacenaron a -80 °C en espera de ser procesados.

Para la purificación de la lectina ANL, los precipitados celulares se resuspendieron en tampón de lisis TSIA-A (20 mM Tris-HCl pH 8,0, 100 mM NaCl, 10 mM imidazol pH 8,0 y 0,04% (p/v) NaN<sub>3</sub>) en presencia de un cóctel de inhibidores de proteasas (Roche). Posteriormente, se lisaron las células usando un sonicador VCX 500 (Sonics), que estaba programado para ciclos de 10 segundos de ultrasonidos durante 5 minutos alternados con 10 segundos de reposo en hielo para evitar aumentos de temperatura. El lisado celular, a continuación, se centrifugó a 40.000 × g durante 30 minutos a 4 °C en un rotor SS34 (SORVALL Evolution RC, Kendro). La fracción soluble se aplicó en una columna HisTrap HP de 5 mL (GE Healthcare Life Sciences) que permite llevar a cabo una cromatografía de afinidad a Ni<sup>2+</sup> inmovilizado sobre una matriz de agarosa. La elución de la proteína retenida se llevó a cabo con un gradiente de imidazol pH 8,0 entre las soluciones TSIA-A y TSIA-B (20 mM Tris-HCl pH 8,0, 100 mM NaCl, 500 mM imidazol pH 8,0 y 0,04% (p/v) NaN<sub>3</sub>). El análisis de las fracciones eluidas se hizo mediante SDS-PAGE y la concentración de proteína se determinó usando el coeficiente de absorción  $E^{0,1\%}$  (1 cm, 280 nm) de 1,685, calculado con el programa ProtParam del servidor ExPASy.

### 3. Técnicas bioquímicas

#### 3.1. Diseño racional de mutantes de LSL<sub>150</sub> mediante mutagénesis dirigida

El diseño racional de mutantes del módulo lectina N-terminal de LSLa se llevó a cabo mediante métodos de mutagénesis dirigida. Se diseñaron un total de siete mutantes que presentan cambios en aminoácidos previsiblemente implicados en la interacción con azúcares. Entre estos mutantes, hay cinco puntuales:  $\alpha 1$  (Ser29His),  $\alpha 2$  (Ser37His),  $\alpha 3$  (Ala43Phe),  $\beta 1$  (Pro77His) y  $\beta 2$  (Ile85His); y dos mutantes triples:  $\alpha F$  (Ser29His, Ser37His y Ala43Phe) y  $\alpha Y$  (Ser29His, Ser37His y Ala43Tyr).

Número (Sentido)	Mutación de LSL <sub>150</sub>	Vector de clonación	Secuencia 5'→3'
M1 (D)	S29H ( $\alpha 1$ )	pKLSL <sub>150</sub>	CGCGCGCAACCATCCCTCTCCCG
M2 (R)	S29H ( $\alpha 1$ )	pKLSL <sub>150</sub>	CGGGAGAGGGATGGTTGCGCGCG
M3 (D)	S37H ( $\alpha 2$ )	pKLSL <sub>150</sub>	GTTGGTCTGCATCCGGTCA
M4 (R)	S37H ( $\alpha 2$ )	pKLSL <sub>150</sub>	TGACCGGATGCAGACCAAC
M5 (D)	A43F ( $\alpha 3$ )	pKLSL <sub>150</sub>	CAACGACCAGTTTACCGACCAG
M6 (R)	A43F ( $\alpha 3$ )	pKLSL <sub>150</sub>	CTGGTCGGTAAACTGGTCGTTG
M7 (D)	P77H ( $\beta 1$ )	pKLSL <sub>150</sub>	CTCGCGTAGGCATGCGGAACC
M8 (R)	P77H ( $\beta 1$ )	pKLSL <sub>150</sub>	GGTTCCGCATGCCTACGCGAG
M9 (D)	I85H ( $\beta 2$ )	pKLSL <sub>150</sub>	GGGCCAACATGATGGCGAC
M10 (R)	I85H ( $\beta 2$ )	pKLSL <sub>150</sub>	GTCGCCATCATGTTGGCCC
M11 (D)	A43Y ( $\alpha Y$ )	pKLSL <sub>150</sub>	CAACGACCAGTATACCGACCAG
M12 (R)	A43Y ( $\alpha Y$ )	pKLSL <sub>150</sub>	CTGGTCGGTATACTGGTCGTTG

**Tabla 3.2.** Secuencias de los oligonucleótidos utilizados para la generación de los diferentes mutantes del módulo lectina LSL<sub>150</sub>. Los sitios donde se inserta la mutación aparecen *subrayados*. Las letras D y R indican el sentido de lectura del oligonucleótido: D significa sentido directo (5'→3') y R, sentido reverso (3'→5' de la hebra codificante).

Para introducir las mutaciones, se partió del plásmido pKLSL<sub>150</sub> y de los respectivos pares de oligonucleótidos diseñados para tal propósito (**Tabla 3.2**). El protocolo de mutagénesis consistió en una reacción de PCR en la que el tiempo de elongación depende del tamaño del vector usado como molde. Después de la reacción de amplificación, las muestras se trataron con 1  $\mu$ L de la endonucleasa *DpnI* (New England Biolabs) durante, al menos,

2 horas a 37 °C. Esta enzima degrada selectivamente el ADN metilado, consiguiéndose eliminar, por tanto, el plásmido molde y dejar sólo las moléculas de plásmido con las mutaciones introducidas resultantes de la amplificación *in vitro*. Tras la digestión enzimática, se transformaron células competentes de *E. coli* DH5 $\alpha$  con 5-10  $\mu$ L de la mezcla de PCR de cada uno de los mutantes. Con las colonias obtenidas se hizo la selección de los clones que incorporaron el plásmido con la mutación introducida correctamente.

Los mutantes fueron expresados y purificados siguiendo el mismo protocolo que el utilizado para la producción y purificación de LSL<sub>150</sub> (ver sección 2.1 de *Materiales y Métodos*).

### **3.2. Análisis de las secuencias de ADN**

El análisis final de las clonaciones y de las mutaciones se realizó mediante secuenciación de ADN por parte del Servicio de Secuenciación y Diagnóstico Molecular (SECUGEN) del Centro de Investigaciones Biológicas (CIB, CSIC). El servicio dispone de un secuenciador automático modelo 3730 DNA Analyzer (Applied Biosystems) y del kit de secuenciación BigDye Terminator v3.1 Cycle Sequencing (Applied Biosystems). Todos los clones se secuenciaron haciendo uso del oligonucleótido *T7 Forward*, que se une a la región promotora presente en todos los plásmidos utilizados en este trabajo y que permite leer en sentido corriente abajo (5'→3').

### **3.3. Espectrometría de masas**

La masa molecular y la homogeneidad de las proteínas se analizó mediante espectrometría de masas en el Servicio de Espectrometría de Masas del Instituto de Química Física "Rocasolano" (IQFR, CSIC). Se empleó un espectrómetro Voyager-DE PRO (Applied Biosystems) equipado con un láser de nitrógeno ( $\lambda = 337$  nm, anchura de pulso 10 ns y frecuencia de 3 Hz) y una fuente iónica con extracción retardada. La ionización de las muestras se llevó a cabo por MALDI (ionización y desorción por láser asistida por matriz) (Hillenkamp *et al.*, 1991). En esta técnica, la muestra es preincubada con una matriz constituida por ácido si-napínico (10 mg/mL en 0,1% TFA:ACN en una relación 70:30 (v:v)). La mezcla sólida resultante se irradia con un láser pulsante cuya energía es absorbida por la matriz, lo que provoca su ionización y la de las moléculas objeto de estudio. Los iones positivos generados por la desorción láser son introducidos en un analizador de tiempo de vuelo (TOF) cuyo tubo tiene 1,3 m de longitud. Aquí los iones son acelerados con un voltaje de 25 kV, deján-

dolos volar libremente hacia el detector. La determinación de las masas se llevó a cabo usando un calibrado externo, constituido por enolasa y anhidrasa carbónica para masas entre 29 y 46 kDa; y por apomioglobina de caballo y tiorredoxina de *E. coli* para el intervalo entre 11 y 29 kDa. Para la obtención de los diferentes espectros se trabajó siempre con muestras de 15-20  $\mu$ L cuya concentración aproximada era 10  $\mu$ M.

### 3.4. Análisis proteómico para la identificación de proteínas

El análisis proteómico para la identificación de algunas de las lectinas purificadas en este trabajo se realizó tanto en la Unidad de Proteómica del Centro Nacional de Investigaciones Cardiovasculares (CNIC, Instituto de Salud Carlos III) como en el Servicio de Proteómica y Genómica del Centro de Investigaciones Biológicas (CIB, CSIC). En ambos casos, las bandas de los geles de poliacrilamida que se quisieron analizar fueron sometidas a una digestión controlada con proteasas. Los péptidos así generados, que constituyen la huella peptídica de la proteína, fueron posteriormente analizados. El análisis de masas de las muestras se realizó de forma automática en un espectrómetro de masas MALDI-TOF/TOF (Bruker) (Suckau *et al.*, 2003), empleando calibración interna. Los datos obtenidos se estudiaron en detalle mediante el programa BioTools 3.0 (Bruker Daltonik) que permite identificar de qué proteína se trata al enfrentar la huella peptídica experimental frente a las bases de datos no redundantes NCBI nr y SwissProt, empleando el programa MASCOT (Matrix Science, Londres).

### 3.5. Técnicas electroforéticas

#### 3.5.1. Electroforesis en geles de agarosa

La electroforesis en geles de agarosa se utilizó para visualizar los fragmentos de ADN amplificados por PCR y para chequear la pureza del ADN plasmídico. La agarosa se disolvió en tampón TAE (40 mM Tris-HCl, 20 mM ácido acético y 2mM EDTA pH 8,0) y su concentración se ajustó en función del tamaño del fragmento que se quería visualizar (Sambrook y Russell, 2001). En nuestro caso, los geles contenían el agente fluorescente GelRed (Biotium) que tiñe selectivamente los ácidos nucleicos. De esta manera, el ADN puede ser visualizado bajo luz ultravioleta con ayuda de un transiluminador. Se usaron los patrones HyperLadderI (Bioline) como referencia para determinar el tamaño de las bandas.

### 3.5.2. Electroforesis en geles de poliacrilamida en presencia de dodecilsulfato sódico (SDS-PAGE)

La electroforesis en geles de poliacrilamida en presencia de dodecilsulfato sódico (SDS-PAGE) se llevó a cabo según el método descrito por Laemmli (Laemmli, 1970), utilizando un sistema Mini-Protean Tetra Cell (Bio-Rad) asociado a una fuente de alimentación que permite variar la diferencia de potencial e intensidad (Bio-Rad). Todas las muestras se prepararon en tampón LB (5 mM Tris-HCl pH 6,8, 2% (p/v) SDS, 10% (v/v) glicerol y 0,005% (p/v) azul de bromofenol) en condiciones reductoras (5% (v/v)  $\beta$ -mercaptoetanol), incubándose durante 5 minutos a 95 °C antes de ser aplicadas. Las electroforesis se desarrollaron siempre a un voltaje constante de 200 V durante 40-50 minutos. Para la estimación de la masa molecular de las muestras se utilizaron los patrones de masa molecular Broad Range (SDS-PAGE Molecular Weight Standards, Bio-Rad). Las bandas de proteína se detectaron mediante tinción con azul brillante de Coomassie R250 al 0,25% (p/v) en metanol:ácido acético:agua (45:5:50 (v:v:v)).

### 3.6. Western blot

La detección de proteínas mediante interacciones específicas de tipo proteína-anticuerpo y/o proteína-lectina se llevó a cabo sobre membranas de nitrocelulosa. Tras una electroforesis de tipo SDS-PAGE, las bandas de proteína fueron transferidas desde el gel de poliacrilamida a una membrana de nitrocelulosa (GE Healthcare Life Sciences). La transferencia se desarrolla en un sistema en el que el gel y la membrana quedan completamente sumergidos en tampón de transferencia (Burnette, 1981). Este sistema permite aprovechar la misma cubeta donde se realiza la electroforesis para llevar a cabo la transferencia (Mini Trans-Blot cell, Bio-Rad). Durante este proceso, la cubeta se acopla a una fuente de alimentación que trabaja a una intensidad constante de 400 mA durante 50-60 minutos. Para evitar sobrecalentamientos, se coloca una unidad de refrigeración junto al sistema de transferencia. Una vez finalizada ésta, la membrana se tiñe primero con rojo Ponceau (0,1% (p/v) Ponceau S en 5% (v/v) ácido acético) y luego se lava con agua destilada para comprobar que las bandas del gel se han transferido correctamente.

La detección posterior de las bandas de interés se puede llevar a cabo con distintos sistemas de revelado. En el caso de la determinación de proteína con el sistema en tándem de dos anticuerpos, el protocolo seguido comienza con el bloqueo de las membranas con 2% de leche en PBS (137 mM NaCl, 2,7 mM KCl, 8 mM Na<sub>2</sub>HPO<sub>4</sub>, 1,46 mM KH<sub>2</sub>PO<sub>4</sub> a pH 7,4)

(PBS-leche) durante 1 hora a temperatura ambiente. Después de lavar las membranas con PBS para eliminar el exceso de agente bloqueante, se dejan incubando con el anticuerpo primario en PBS-leche a 4 °C durante toda la noche y con agitación constante. Seguidamente, se lavan de nuevo con PBS y se vuelven a incubar con el anticuerpo secundario conjugado con HRP (peroxidasa de rábano) en PBS-leche durante 1 hora a temperatura ambiente. Finalmente, las membranas se lavan exhaustivamente con PBS y se revelan utilizando el kit SIGMAFAST DAB with Metal Enhancer (Sigma-Aldrich). El DAB (3,3'-diaminobenzidina) actúa como sustrato de la HRP generando un producto insoluble y coloreado que tiñe la membrana.

Por otro lado, cuando la detección es dependiente de la interacción proteína-lectina, el protocolo seguido fue ligeramente diferente. Esta variante de *western blot* que utiliza lectinas en el reconocimiento se denomina con el nombre de *lectin blot*. En primer lugar, se bloquean las membranas con TBSB (50 mM Tris-HCl pH 7,8, 150 mM NaCl y 0,25% Tween-20) durante 1 hora a temperatura ambiente. Tras esta etapa, se lavan con TBS (50 mM Tris-HCl pH 7,8 y 150 mM NaCl) y se incuban con las lectinas biotiniladas (10 µg/mL) en TBSI (TBS con Tween-20 al 0,1%) durante 90 minutos a temperatura ambiente. Después se lavan de nuevo con TBS y se dejan incubando con estreptavidina-HRP (1 µg/mL) en TBSI durante 35 minutos a temperatura ambiente. Por último, las membranas se someten a varios lavados con TBS y se revelan con el sistema SIGMAFAST DAB with Metal Enhancer (Sigma-Aldrich). En este caso, las membranas se escanearon con un sistema de detección de imagen ChemiDoc XRS System (Bio-Rad) y se analizaron con ayuda del software Quantity One (Bio-Rad).

### 3.7. Producción de anticuerpos policlonales anti-lectina

Para facilitar la identificación mediante técnicas inmunoquímicas de las proteínas LSLa, LBL o de proteínas de fusión que incluyan los módulos lectina LSL<sub>150</sub> o LBL<sub>152</sub>, se produjeron anticuerpos policlonales específicos de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub>. Los anticuerpos se obtuvieron en el animalario del Centro de Investigaciones Biológicas (CIB, CSIC). La producción de anticuerpos se llevó a cabo en sendos conejos, inoculándoles los módulos lectina (1 mg/mL) en PBS (1 mL por inyección). Después de la inmunización, la sangre de los animales se centrifugó a 3.000 × g durante 15 minutos a 4 °C (IEC CL40 Centrifuge, Thermo Scientific) con la finalidad de extraer el suero, que contiene los correspondientes anticuerpos.



La especificidad de los anticuerpos se comprobó mediante *western blot* frente a un set de lectinas con plegamiento tipo trébol  $\beta$ : LSL<sub>150</sub>, LBL<sub>152</sub>, viscumina de *Viscum album* (VAA), ricina y aglutinina de *Ricinus communis* (RCA). Los resultados demostraron que los anticuerpos anti-LSL<sub>150</sub> son específicos del módulo lectina de *L. sulphureus*; mientras que los anticuerpos anti-LBL<sub>152</sub>, tienen la capacidad de reconocer tanto LBL<sub>152</sub> como LSL<sub>150</sub>.

### 3.8. Cromatografía analítica de exclusión molecular

La cromatografía de exclusión molecular o por tamaño es una técnica cromatográfica que permite separar moléculas en función de su tamaño y relación forma/tamaño. La columna está constituida por una matriz sólida o soporte compuesto por micropartículas porosas formadas por enrejados de fibras. Asumiendo estructuras globulares, las proteínas cuya masa molecular es mayor, previsiblemente van a eluir antes de la columna debido a que son excluidas del entramado que el soporte presenta. Por el contrario, las moléculas más pequeñas verán retardada su elución. Con ciertas limitaciones, esta metodología es muy útil a la hora de estimar la masa molecular de proteínas en solución y por ello estimar su estado oligomérico.

Las cromatografías se realizaron en un sistema BioLogic DuoFlow FPLC (Bio-Rad) equipado con una columna Superdex 200 10/300 GL (GE Healthcare Life Sciences). En primer lugar, la columna se equilibró con tampón TSA para, posteriormente, aplicar la proteína (~100  $\mu$ g) e iniciar el método de elución a un flujo de 0,5 mL/min.

La columna fue calibrada con una mezcla de proteínas de masa molecular conocida (Gel Filtration Standard, Bio-Rad): tiroglobulina (670 kDa),  $\gamma$ -globulina (158 kDa), ovoalbúmina (44 kDa), mioglobina (17 kDa) y vitamina B<sub>12</sub> (1,35 kDa). Teniendo en cuenta el volumen en el que eluyen estas proteínas patrones, es posible estimar la masa molecular de la muestra en cuestión mediante interpolación en la recta previamente determinada que relaciona el logaritmo de la masa molecular de las proteínas estándar con sus coeficientes de reparto ( $K_{av}$ ) (**Ecuaciones 3.1 y 3.2**), donde  $V_e$  es el volumen de elución,  $V_0$  es el volumen de exclusión y  $V_t$  es el volumen total de la columna.

$$K_{av} = \frac{V_e - V_0}{V_t - V_0} \quad \text{(Ecuación 3.1)}$$

$$\log MW = aK_{av} + b \quad \text{(Ecuación 3.2)}$$

### 3.9. Ensayos de actividad enzimática

Se llevaron a cabo ensayos de actividad enzimática con la  $\beta$ -galactosidasa (Bgl) de *Lactobacillus plantarum* CECT 748T y la lipasa BTL2 de *Geobacillus thermocatenolatus*.

La actividad glicosidasa de la enzima se midió a partir de la hidrólisis frente a los siguientes sustratos: *p*-nitrofenil- $\beta$ -D-glucósido (*p*NPG), *p*-nitrofenil- $\beta$ -D-galactósido (*p*NPGal), *o*-nitrofenil- $\beta$ -D-galactósido (*o*NPG) y *o*-nitrofenil- $\beta$ -D-galactósido-6-fosfato (*o*NPG6P). La aparición de coloración amarilla indica que la reacción de hidrólisis transcurre con éxito y que, por tanto, existe actividad glicolítica. El volumen de reacción se ajustó a 1 mL y se llevó a cabo en 50 mM tampón fosfato pH 7,0 en presencia de 5 mM de sustrato. La reacción se realizó a 37 °C y se siguió a diferentes intervalos de tiempo. Además, se estudió el efecto del agente reductor 1,4-ditiotreitol (DTT) sobre la capacidad hidrolítica de la enzima.

Por otro lado, la actividad lipasa se registró midiendo el incremento de absorbancia a 348 nm producido por la liberación de *p*-nitrofenol durante la hidrólisis de 0,4 mM butirato de *p*-nitrofenilo (*p*NPB) en 25 mM fosfato de sodio pH 7,0 a 25 °C, usando un baño termostático con agitación magnética. La reacción enzimática tiene lugar al añadir 50  $\mu$ L de enzima sobre 2,5 mL de solución con el sustrato. La unidad internacional de actividad *p*NPB (U) se define como la cantidad de enzima necesaria para hidrolizar 1  $\mu$ mol de *p*NPB/min bajo las condiciones anteriormente descritas (Fernández-Lorente *et al.*, 2008).

## 4. Técnicas de caracterización biofísica de proteínas

### 4.1. Espectroscopía de fluorescencia en estado estacionario

La espectroscopía de fluorescencia es una técnica muy útil en estudios de cambios conformacionales en macromoléculas biológicas por su elevada sensibilidad, ya que responde a cambios muy sutiles del microentorno de los fluoróforos<sup>14</sup>. Las proteínas presentan fluorescencia intrínseca en la región del UV próximo debido a la presencia de residuos de triptófano y tirosina esencialmente. En algunos casos es posible unir a la molécula de interés un fluoróforo extrínseco con el fin de dotar a la molécula de propiedades espectrales adecuadas para un experimento dado.

La fluorescencia es un fenómeno de relajación radiativa que tiene lugar cuando se produce una transición electrónica desde el modo vibracional de menor energía del estado excitado (singlete) a un modo vibracional del estado fundamental, emitiéndose un fotón de

luz. Este fenómeno físico complejo es utilizado para caracterizar sistemas biológicos, determinando los respectivos espectros de emisión y excitación y estudiando ciertos aspectos como la anisotropía de fluorescencia<sup>3</sup> en estado estacionario.

#### 4.1.1. Espectros de emisión y excitación de fluorescencia

A pesar de que las proteínas tienen diferentes fluoróforos naturales, los espectros de emisión fluorescente de una proteína están normalmente dominados por la emisión de tirosinas (Tyr) y triptófanos (Trp) debido a su mayor sensibilidad (García-Segura *et al.*, 2002). El triptófano, de hecho, es el fluoróforo intrínseco por excelencia de las proteínas. Presenta un mecanismo fotofísico complejo que, aunque limita en muchos casos su utilización como sonda fluorescente, permite una caracterización preliminar de proteínas no marcadas.

Los espectros de fluorescencia se realizaron en un espectrofluorímetro PC1 (ISS Champaign, Illinois) en el modo de conteo de fotones. Como fuente de luz se utilizó una lámpara de xenón de alta presión Osram (450 W) que emite de forma continua desde 250 nm hasta el infrarrojo, siendo seleccionada la longitud de onda de excitación por medio de un doble monocromador. La intensidad de fluorescencia se midió a 90° con respecto del haz de excitación, empleándose cubetas de cuarzo (Starna) de 0,2 cm de paso óptico de excitación y 1 cm de emisión. La temperatura del compartimento de las muestras se mantuvo a 25 °C, controlada con un Peltier.

El espectro de emisión de fluorescencia se obtuvo registrando la intensidad de la radiación emitida por la muestra en un intervalo de longitudes de onda tras la excitación a una longitud de onda seleccionada. Del mismo modo, la variación en la intensidad de fluorescencia en función de la longitud de onda de excitación para una longitud de emisión fija constituye el espectro de excitación de fluorescencia.

Incluso en ausencia de polarizadores, la radiación de excitación está parcialmente polarizada debido al monocromador de excitación. Esto da lugar a que la emisión pueda estar parcialmente polarizada en función de la vida media del estado excitado. Esta polarización de fluorescencia puede constituir una fuente significativa de errores sistemáticos en las medidas de intensidad de fluorescencia si el sistema de detección presenta diferente sensibilidad hacia los componentes vertical y horizontal. Esta distorsión debida al efecto de la polarización de la muestra se ha eliminado experimentalmente excitando a las muestras con luz verticalmente polarizada y midiendo la intensidad de fluorescencia polarizada a

54,7° (el llamado ángulo mágico) de la vertical, siendo la señal medida de esta manera proporcional a la intensidad de fluorescencia total (Spencer y Weber, 1970).

#### 4.1.2. Anisotropía de fluorescencia en estado estacionario

Cuando se excita un fluoróforo con luz polarizada plana, se produce un fenómeno de fotoselección en tanto que sólo aquellas moléculas convenientemente orientadas absorberán fotones. Durante el tiempo en que la molécula se encuentra en el estado excitado, ésta puede rotar libremente, dando lugar a que los fotones emitidos tengan planos de oscilación distintos de aquel de la luz incidente. Si la luz emitida se analiza con un polarizador, de tal manera que únicamente deja pasar luz emitida en el mismo plano que el de la luz incidente (sea  $I_{\parallel}$  esta intensidad), sólo se registrará una fracción de la luz total emitida por fluorescencia. Por el contrario, si el polarizador de emisión se orienta perpendicularmente al de excitación, se detectará el componente perpendicular de los fotones totales emitidos ( $I_{\perp}$ ). Para evaluar la capacidad que tiene la muestra de emitir luz en un plano u otro, se define la magnitud anisotropía de fluorescencia ( $r$ ) (**Ecuación 3.3**) (Jablonski, 1957). Su medida proporciona información sobre la difusión rotacional del estado excitado, es decir, sobre las propiedades hidrodinámicas del fluoróforo.

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \quad (\text{Ecuación 3.3})$$

Las medidas de anisotropía de fluorescencia en estado estacionario se realizaron en el un espectrofluorímetro PC1 (ISS Champaign, Illinois), al que se le incorporaron polarizadores Glan-Thompson en los canales de emisión y excitación de fluorescencia. Las muestras fueron excitadas con luz polarizada verticalmente y se midieron las componentes de la intensidad de fluorescencia polarizada en las direcciones paralela ( $I_{\parallel}$ ) y perpendicular ( $I_{\perp}$ ) al plano de polarización de la luz de excitación.

Para el cálculo de la anisotropía de fluorescencia se utilizó la **Ecuación 3.4**. Esta ecuación difiere de la teórica (**Ecuación 3.3**) por la introducción del factor de corrección G. Debe hacerse esta corrección debido a la diferente sensibilidad del canal de emisión a la luz polarizada horizontal y verticalmente. Este factor es función de la longitud de onda de los fotones emitidos y se determina midiendo la respuesta del detector utilizando radiación de excitación polarizada horizontalmente y colocando secuencialmente el polarizador de emisión en las posiciones vertical y horizontal. Cuando la muestra se excita con luz polarizada

horizontalmente, las diferencias entre la intensidad medida con polarización vertical y horizontal en la emisión se deben únicamente a la diferencia de sensibilidad del sistema de detección a la luz polarizada. El valor de  $G$  se determinó mediante la expresión de la **Ecuación 3.5**, dependiente de  $I_{HV}$  e  $I_{HH}$ , cuyos subíndices se refieren a las orientaciones de los polarizadores cuando la excitación está polarizada horizontalmente.

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \quad (\text{Ecuación 3.4})$$

$$G = \frac{I_{HV}}{I_{HH}} \quad (\text{Ecuación 3.5})$$

## 4.2. Dicroísmo circular

El dicroísmo circular (DC) es una técnica espectroscópica que proporciona información sobre la estructura secundaria y terciaria de proteínas, permitiendo a su vez registrar cambios en dicha estructura de un modo sencillo.

El espectro de DC de una proteína en la región del UV lejano (180-250 nm) es el resultado de la contribución de los diferentes elementos de estructura secundaria (hélices  $\alpha$ , láminas  $\beta$ , giros  $\beta$  y otras conformaciones como la estructura desordenada) que la proteína presenta (Brahms y Brahms, 1980), cada uno de los cuales contribuye al espectro total con una señal característica. El enlace peptídico es el principal cromóforo en esta región y su disposición espacial afecta al comportamiento espectroscópico, razón por la cual podemos extraer información de carácter estructural.

En la región del UV cercano (250-350 nm) los cromóforos más importantes son los aminoácidos aromáticos (triptófano, tirosina y fenilalanina) y los enlaces disulfuro (Strickland, 1974; Kahn, 1979). Dado que la señal de estos cromóforos depende de su entorno, los espectros en esta región son un reflejo de la conformación global de la proteína (estructura terciaria).

Los espectros de DC se registraron en un espectropolarímetro J-810 (Jasco) con un sistema acoplado de Peltier para el control de la temperatura. Se utilizó un tiempo de respuesta de 2 segundos, una velocidad de barrido de 50 nm/min y los datos fueron recogidos cada 0,2 nm. Para las medidas en el UV lejano se utilizó una concentración de proteína de 0,2-0,3 mg/mL y cubetas de cuarzo con un paso óptico de 0,1 cm. Por el contrario, en la región del UV cercano, la concentración de proteína empleada fue de 1 mg/mL y el paso óptico de las cubetas de 1 cm. Para cada espectro se recogieron cuatro acumulaciones con

el fin de obtener una buena relación señal-ruido. La recogida de datos experimentales y el análisis preliminar de los mismos se realizó con el programa *Spectra Manager* (Jasco). Antes de medir cada muestra, se recogió un espectro del tampón que servía como blanco. Los datos experimentales, en unidades de miligrados, se transformaron a unidades de elipticidad molar por residuo ( $[\theta]_{\text{MRW}}$ ) ( $\text{grado cm}^2 \text{ dmol}^{-1}$ ).

La estabilidad térmica de las lectinas se investigó por DC midiendo variaciones en la elipticidad de la proteína a una longitud de onda determinada inducidas por un aumento progresivo de la temperatura. Los experimentos se realizaron con una elevación constante de la temperatura de  $40 \text{ }^\circ\text{C/h}$ , registrándose los valores de elipticidad cada  $0,2 \text{ }^\circ\text{C}$ . Los perfiles de desnaturalización térmica a longitud de onda fija se ajustaron a una o dos sigmoideas, dependiendo del número de transiciones observadas, haciendo uso de la **Ecuación 3.6**, donde  $[\theta](T)$  es la función elipticidad molar en función de la temperatura,  $[\theta]_N$  es la elipticidad molar de la proteína en su estado nativo a una longitud de onda dada,  $n$  es el número de transiciones,  $\Delta[\theta]_i$  es la variación de la elipticidad asociada a la transición  $i$ ,  $R$  es la constante de los gases ideales, y  $T_{mi}$  y  $HD_i$  son la temperatura media de transición y el parámetro que describe la cooperatividad (pendiente) de la transición correspondiente, respectivamente (Campanero-Rhodes *et al.*, 2005).

$$[\theta](T) = [\theta]_N + \sum_{i=1}^n \frac{\Delta[\theta]_i e^{\frac{-HD_i(T_{mi}-T)}{RT_{mi}T}}}{1 + e^{\frac{-HD_i(T_{mi}-T)}{RT_{mi}T}}} \quad (\text{Ecuación 3.6})$$

### 4.3. Calorimetría isotérmica de titulación (ITC)

La calorimetría isotérmica de titulación (ITC, del inglés *Isothermal Titration Calorimetry*) permite estudiar a temperatura constante la energética de una reacción química iniciada con la mezcla de dos componentes (formación de complejos, catálisis enzimática, etc.), midiendo el calor absorbido o producido en el proceso. De hecho, se trata de la técnica experimental más directa para determinar los parámetros termodinámicos asociados a la formación de complejos proteína-ligando (Wiseman *et al.*, 1989; Velazquez-Campoy *et al.*, 2004).

Un calorímetro de titulación isotérmica consta de dos células idénticas, una de las cuales se utiliza como célula de muestra y la otra como referencia. En un experimento de ITC, la reacción se estudia midiendo el calor producido o absorbido al inyectar, en pasos

sucesivos, uno de los componentes (generalmente, el ligando) en la solución del otro (generalmente, la proteína) situado en la celda de muestra. El calor absorbido o producido tras cada inyección genera una diferencia de temperatura entre ambas células que se compensa automáticamente mediante un sistema de retroalimentación, obteniéndose una serie de picos cuya integral corresponde al calor producido o absorbido en cada inyección como consecuencia de la reacción. En la formación de complejos, a medida que aumenta el grado de saturación de la proteína por el ligando, la intensidad de los picos disminuye hasta alcanzar un valor constante que corresponde al calor de dilución del ligando. Este calor de dilución se cuantifica mediante valoraciones en paralelo, en las que la disolución del ligando se inyecta sobre la celda de reacción conteniendo el tampón de medida en el que se haya la proteína, y es descontado del calor total para obtener el calor debido a la interacción.

El calor de interacción en cada inyección está relacionado con la variación de entalpía del proceso estudiado y el grado de saturación alcanzado, lo que permite obtener los parámetros de la interacción [ $n$  (número de sitios),  $\Delta H$  (entalpía de unión) y  $K_a$  (constante de asociación)] para cada clase de sitios de unión a partir del ajuste de los datos experimentales (calor por mol de ligando inyectado vs relación molar ligando/proteína), utilizando modelos apropiados. El valor de la energía de Gibbs ( $\Delta G$ ) y de la entropía de interacción ( $\Delta S$ ) se determina mediante las expresiones de las **Ecuaciones 3.7** y **3.8**, resultado de la aplicación de las leyes fundamentales de la termodinámica.

$$\Delta G = -RT \ln K_a \quad \text{(Ecuación 3.7)}$$

$$\Delta G = \Delta H - T\Delta S \quad \text{(Ecuación 3.8)}$$

La preparación de las muestras para ITC requirió una diálisis previa muy exhaustiva de las lectinas frente al tampón sin ligando. Después de la diálisis, la concentración final de las proteínas se ajustó a un valor próximo a los 250  $\mu\text{M}$ . El tampón procedente de la última diálisis fue utilizado para la preparación del ligando. El ligando, a una concentración inicial de 30-60 mM, fue añadido a la disolución de proteína, presente en la celda de reacción, en una serie de inyecciones cuyo volumen y número total fue ajustado en función del experimento. Los experimentos se llevaron a cabo en un microcalorímetro VP-ITC (GE Healthcare Life Sciences) y la recogida y el análisis de datos se realizaron con VPViewer y MicroCal, LLC ITC (GE Healthcare Life Sciences), respectivamente.

#### 4.4. Calorimetría diferencial de barrido (DSC)

La calorimetría diferencial de barrido (DSC del inglés *Differential Scanning Calorimetry*) es una técnica usada ampliamente para medir la estabilidad de proteínas, ácidos nucleicos y membranas biológicas (Johnson, 2013). Permite caracterizar de forma directa la capacidad calorífica del sistema estudiado como una función continua de la temperatura a presión constante ( $C_p$ ). El experimento tiene lugar en condiciones adiabáticas y el calorímetro consta de dos celdas (muestra y referencia), que se mantienen a una diferencia de temperatura constante y casi nula mediante un sistema de retroalimentación que minimiza la diferencia de temperatura entre ambos cuando su capacidad calorífica difiere debido a los cambios conformacionales inducidos por la temperatura en la biomolécula. La energía empleada para igualar la temperatura de la celda de muestra y referencia es proporcional a los cambios de capacidad calorífica de la muestra, y su variación con la temperatura (termograma) se caracteriza por la presencia de uno o más picos de absorción de calor que son reflejo de los cambios conformacionales que está sufriendo la muestra. En el caso particular de las proteínas, por ejemplo, la pérdida de su estructura terciaria nativa es un proceso globalmente endotérmico que aparece con, al menos, un máximo en el termograma. A partir de estos resultados, se puede determinar la entalpía asociada al proceso que ha tenido lugar, mediante la **Ecuación 3.9**, donde  $T_1$  y  $T_2$  son, respectivamente, las temperaturas a las cuales el proceso comienza y termina, respectivamente, y  $C_{p,ex}$  es la capacidad calorífica en exceso debida a la energía transferida para igualar la temperatura de las celdas.

$$\Delta H = \int_{T_1}^{T_2} C_{p,ex} dT \quad (\text{Ecuación 3.9})$$

En condiciones de equilibrio, la relación existente entre  $C_{p,ex}$ ,  $\Delta H$  y  $\Delta S$  permite realizar una caracterización termodinámica completa del proceso de desnaturalización (número de estados, mecanismo de reacción, etc.); mientras que en procesos irreversibles (muy frecuentes en la desnaturalización de proteínas), el acceso a la información termodinámica y/o estructural está limitado a aquellos casos en que la población de formas reversiblemente desnaturalizadas sigue siendo significativa en el intervalo de temperaturas en que tiene lugar la transición o transiciones.

Los experimentos de DSC se llevaron a cabo en un microcalorímetro VP-DSC (MicroCal). Las muestras se prepararon a una concentración de 0,3 mg/mL de proteína tanto en tampón TSA como en tampón TSAL (previa diálisis) para ver el efecto de la lactosa sobre



la estabilidad de la proteína. Los perfiles de estabilidad térmica se obtuvieron en el rango de 25 a 95 °C, utilizando una rampa de temperatura de 80 °C/h.

#### 4.5. Ultracentrifugación analítica

La ultracentrifugación analítica permite determinar la masa molecular, forma global aproximada y grado de homogeneidad de biomoléculas en solución, sin interacción con ninguna matriz o superficie, así como caracterizar procesos de asociación cuando la masa molecular del complejo resultante difiere significativamente de las de los componentes aislados (Rivas *et al.*, 1999). Desde un punto de vista experimental, existen dos aproximaciones básicas para estudiar un sistema biológico: equilibrio de sedimentación y velocidad de sedimentación.

El equilibrio de sedimentación es una técnica en la que las macromoléculas estudiadas son sometidas a la acción de un campo centrífugo moderado ( $\omega < 45.000$  rpm), lo que permite alcanzar un equilibrio entre el proceso de sedimentación de las partículas debido a la acción de la fuerza centrífuga y el de difusión causado por el gradiente de concentración creado por la centrifugación. Esta técnica permite obtener información sobre la masa molecular promedio de las partículas en solución, la homogeneidad de la muestra y las constantes de equilibrio y estequiometría de los complejos de sistemas asociativos (Rivas *et al.*, 1999).

La técnica de velocidad de sedimentación permite analizar el comportamiento hidrodinámico de la partícula que sedimenta, el cual depende de su masa y su forma. En los experimentos de velocidad, las moléculas son sometidas a un campo centrífugo elevado (velocidades  $\omega \geq 45.000$  rpm), formándose un frente en movimiento con un transporte neto de la masa hacia el fondo de la célula, al ser mayor la fuerza de sedimentación que la de difusión. El análisis de los perfiles de sedimentación a intervalos regulares de tiempo permite calcular los coeficientes de sedimentación y difusión, y a partir de ellos analizar o modelar, por ejemplo, la forma hidrodinámica de las proteínas (Lebowitz *et al.*, 2002).

En este trabajo se han utilizado ambas aproximaciones para la caracterización del estado oligomérico tanto de la proteína de fusión LSLt-Bgl como de la proteína uBgl. Para ello se utilizó una ultracentrífuga analítica Optima-XLA (Beckman-Coulter) equipada con un sistema óptico de detección UV-Vis.

Los experimentos de equilibrio de sedimentación se llevaron a cabo en células de doble sector con ventanas de cuarzo de 12 mm de paso óptico y usando un rotor An-50 Ti

(Beckman-Coulter) con un cuerpo central de seis agujeros. Los ensayos se realizaron a 20 °C, utilizando muestras cuya absorbancia a 280 nm se encontraba en el rango 0,2-0,6, y tres velocidades distintas (10.000, 13.000 y 18.000 rpm). Con el volumen de muestra utilizado (80 µL), el equilibrio se alcanzó en 12 horas. La contribución de la absorbancia del tampón se corrigió sedimentando el material proteico a 45.000 rpm durante 6 horas y haciendo la lectura de la absorbancia en las proximidades del menisco. La masa molecular aparente promedio se calculó ajustando los datos experimentales a la ecuación matemática que describe la distribución radial del gradiente de concentración en el equilibrio para una especie ideal, utilizando el programa HeteroAnalysis ([www.biotech.uconn.edu](http://www.biotech.uconn.edu), Analytical Ultracentrifugation Facility, University of Connecticut, EEUU), tras comprobar la conservación de la masa.

Los experimentos de velocidad de sedimentación se llevaron a cabo a 20 °C en células de doble sector (12 mm de paso óptico) con un cuerpo central de dos agujeros, utilizando muestras de proteína con una absorbancia a 280 nm de 0,5 y un volumen de 400 µL, y aplicando una velocidad de 45.000 rpm. Los coeficientes de sedimentación se calcularon a partir del ajuste de los perfiles de sedimentación utilizando el programa SEDFIT y el valor del coeficiente de sedimentación se corrigió a las condiciones estándar [agua a 20 °C ( $s_{20,w}$ )] con el mismo programa. Se utilizó el programa SEDNTERP (Laue, 1997) para calcular el valor del volumen específico parcial de las proteínas, valiéndose de la composición de aminoácidos, y para estimar la densidad y viscosidad de los tampones de medida.

## 5. Análisis de la especificidad de las lectinas

El conocimiento de qué azúcares son reconocidos por las lectinas objeto de estudio brinda una oportunidad para comprender mejor la función biológica de las mismas. Existen multitud de técnicas que se han desarrollado con esta intención (Bertozzi y Kiessling, 2001), pero los análisis masivos miniaturizados o *microarrays* son actualmente los más efectivos (Park *et al.*, 2013). La combinación de los análisis de especificidad de las lectinas con estudios de carácter estructural permite tener una idea más global acerca del posible rol en la naturaleza de estas proteínas (Petukhov *et al.*, 1999).

### 5.1. *Glycan array screening*

Gracias al Consorcio de Glicómica estructural (CFG, del inglés *Corsortium for Functional Glycomics*) se ha llevado a cabo un estudio de la especificidad de los módulos lectina N-

terminales LSL<sub>150</sub> y LBL<sub>152</sub>. Este estudio está basado en la tecnología de las micromatrices de glicanos o *glycan arrays*, una poderosa herramienta que permite identificar qué tipo de azúcares son reconocidos por las proteínas mediante interacciones específicas. Estos ensayos se realizan sobre una matriz que presenta azúcares dispuestos de forma ordenada en una superficie sólida y con una orientación adecuada para que puedan ser reconocidos por las correspondientes lectinas (Fazio *et al.*, 2002; Fukui *et al.*, 2002; Houseman y Mrksich, 2002; Park y Shin, 2002; Wang *et al.*, 2002). La inmovilización de azúcares ayuda a que las proteínas que unen débilmente azúcares en solución puedan unirse fuertemente, permitiendo un análisis rápido de la interacción proteína-azúcar. Algunas de las proteínas de unión a azúcares o tipo GBP (*glycan binding protein*) presentan más de un sitio de unión a ligando, por lo que pueden unirse a varias moléculas de azúcar simultáneamente traduciendo, en la práctica, en una mayor afinidad (Park *et al.*, 2013).

Las muestras de LSL<sub>150</sub> y LBL<sub>152</sub> que se utilizaron para llevar a cabo el *glycan array* se produjeron con una cola de seis histidinas en posición C-terminal. Esta etiqueta de afinidad tuvo que ser añadida *ad hoc* para la posterior etapa de revelado, en la que se utilizan anticuerpos monoclonales anti-*His tag* ( $\alpha$ -His). La producción de las lectinas con la etiqueta de fusión se inició con la clonación de los fragmentos *lsl*<sub>150</sub> y *lbl*<sub>152</sub> mediante fusión por PCR (de las Rivas *et al.*, 2007) en el vector pURI3-Cter (Curiel *et al.*, 2011), utilizando los pares de oligonucleótidos 17/18 y 19/20 (**Tabla 3.3**), respectivamente. Ambas lectinas fueron expresadas tras inducir cultivos crecidos en medio LB-Ap a 37 °C con IPTG (0,4 mM) durante 20 horas a 22 °C. Después de la inducción, las células se recogieron mediante centrifugación para posteriormente ser resuspendidas en 15-20 mL de tampón TSIA-A en presencia de un cóctel de inhibidores de proteasas. Tras la lisis, la purificación de las lectinas a partir de la fracción soluble se llevó a cabo en un solo paso mediante cromatografía de afinidad a Ni<sup>2+</sup> usando una columna HisTrap de 5 mL (GE Healthcare Life Sciences) acoplada a un equipo ÄKTAprime plus (GE Healthcare Life Sciences). Para la elución se preparó un gradiente de imidazol a pH 8,0 usando los tampones TSIA-A y TSIA-B. Las fracciones con proteína se concentraron y se terminaron de purificar en un segundo paso cromatográfico con una columna HiLoad 16/600 Superdex 75 Prep Grade (GE Healthcare Life Sciences) de exclusión por tamaño, que permite eliminar el imidazol y apartar los agregados solubles. Las fracciones finales, ya en tampón TSA, se analizaron mediante SDS-PAGE y la concentración se estimó usando los coeficientes de absorción  $E^{0,1\%}$  (1 cm, 280 nm) siguientes: 1,523 para LSL<sub>150</sub>-His y 1,381 para LBL<sub>152</sub>-His.

Número (Sentido)	Gen que se amplifica	Vector de clonación	Secuencia 5'→3'	Vector
17 (D)	<i>lsl<sub>150</sub>-His</i>	pURI3-Cter	<i>taactttaagaaggagatat</i> <i>acat</i> ATGACCGACATCTACATCCCG	pURI3- LSL <sub>150</sub> His
18 (R)	<i>lsl<sub>150</sub>-His</i>	pURI3-Cter	<i>gctattaatgatgatgatgatgatg</i> ATCCTCGAAGAGGAAGGTGAAG	pURI3- LSL <sub>150</sub> His
19 (D)	<i>lbl<sub>152</sub>-His</i>	pURI3-Cter	<i>taactttaagaaggagatat</i> <i>acat</i> ATGTCAAACGAGTACAATCCCCC	pURI3- LBL <sub>152</sub> His
20 (R)	<i>lbl<sub>152</sub>-His</i>	pURI3-Cter	<i>gctattaatgatgatgatgatgatg</i> GTCGAATAAGATAGTGAAATATTG	pURI3- LBL <sub>152</sub> His

**Tabla 3.3.** Secuencias de los oligonucleótidos utilizados para la construcción de las lectinas con cola de histidinas en posición C-terminal. En *minúscula y cursiva* aparecen los sitios que hibridan directamente con el plásmido para la clonación independiente de enzimas de restricción y ligasa. En *mayúscula* se indican las regiones que hibridan con el inserto. Las letras D y R indican el sentido de lectura del oligonucleótido: D significa sentido directo (5'→3') y R, sentido reverso (3'→5' de la hebra codificante).

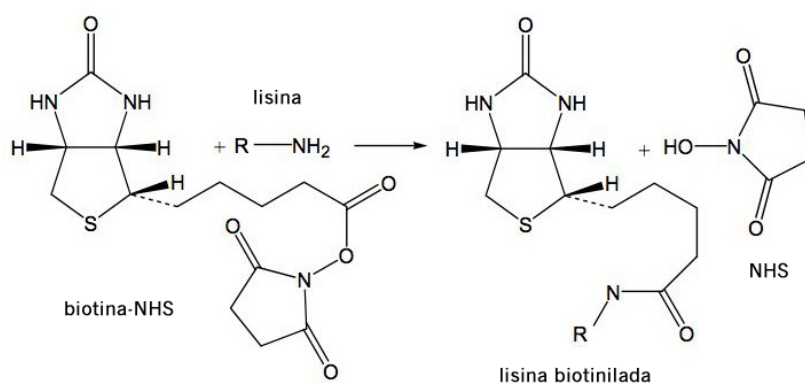
Cada ensayo de unión consistió en incubar los módulos lectina a diferentes concentraciones (0,02, 0,2, 2, 20 y 200 µg/mL) frente a matrices que presentaban 611 azúcares inmovilizados de diferente complejidad (*Mammalian Printed Array Ver. 5.0*). Los azúcares son inmovilizados en una placa de vidrio que previamente es activada con el reactivo químico N-hidroxisuccinimida, que permite establecer un enlace covalente entre la superficie de la placa y azúcares funcionalizados con grupos amino (Blixt *et al.*, 2004). La unión a azúcares fue llevada a cabo incubando cada placa con 70 µL de lectina en tampón TSMBB (20 mM Tris-HCl pH 7,5, 150 mM NaCl, 2mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 0,05% Tween-20 y 1% BSA) durante 1 hora a temperatura ambiente. Seguidamente, las placas se lavaron de forma exhaustiva con las soluciones TSMW (20 mM Tris-HCl pH 7,5, 150 mM NaCl, 2mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub> y 0,05% Tween-20) y TSM (20 mM Tris-HCl pH 7,5, 150 mM NaCl, 2mM CaCl<sub>2</sub> y 2 mM MgCl<sub>2</sub>). Por último, se realizó la incubación con 70 µL del anticuerpo α-His marcado con AlexaFluor-488 (fluoróforo que se utiliza para cuantificar la unión) durante 1 hora a temperatura ambiente. De esta manera y tras nuevos lavados con TSMW, TSM y agua, las placas se introdujeron en el escáner (ProScan Array, Perkin Elmer) donde fueron registradas todas las intensidades correspondientes a la unión de las lectinas a cada uno de los azúcares. Todos los detalles de los protocolos seguidos en la realización del *glycan array* están incluidos en la página web del CFG y descritos por Smith y colaboradores (Smith *et al.*, 2010).

## 5.2. Marcaje con biotina

La modificación química covalente de proteínas con biotina es uno de los marcajes más ampliamente usados en biociencias (Hermanson, 2008). La biotina se utiliza principalmente por tener la capacidad de interaccionar con la estreptavidina de forma rápida y muy estable. Aunque la biotina no da ninguna señal por sí misma, la estreptavidina puede ser marcada con una gran variedad de moléculas como fluoróforos o enzimas. De esta manera, la interacción biotina-estreptavidina se convierte en un método sensible de detección de proteínas inmovilizadas.

La biotinylación de las lectinas se llevó a cabo utilizando el kit Amersham ECL Protein Biotinylation Module (GE Healthcare Life Sciences), que utiliza como reactivo el biotinato de N-hidroxisuccinimilo (biotina-NHS). Este reactivo reacciona específicamente con los grupos  $\epsilon$ -amino de los residuos de lisina (Lys) accesibles en la superficie de las proteínas, formándose un enlace covalente entre el grupo amino y el éster de biotina (**Figura 3.2**).

Para llevar a cabo la reacción de biotinylación, se partió de 200  $\mu$ L de lectina (1 mg/mL) que habían sido dializados previamente frente a tampón TSAL. A cada muestra se le añadieron 10,5  $\mu$ L de tampón bicarbonato 40 mM, pH 8,6 y 10  $\mu$ L de biotina-NHS 2 $\times$ . La mezcla de reacción se dejó 1 hora con agitación a temperatura ambiente. Posteriormente, las lectinas se dializaron exhaustivamente a 4 °C para eliminar completamente la lactosa (se hicieron, al menos, cuatro cambios cada 4 horas frente a 1 litro de solución sin lactosa). Finalmente, las lectinas biotiniladas se desecaron con un concentrador Savant SpeedVac (Thermo Scientific) y se almacenaron a 4 °C en espera de ser utilizadas.



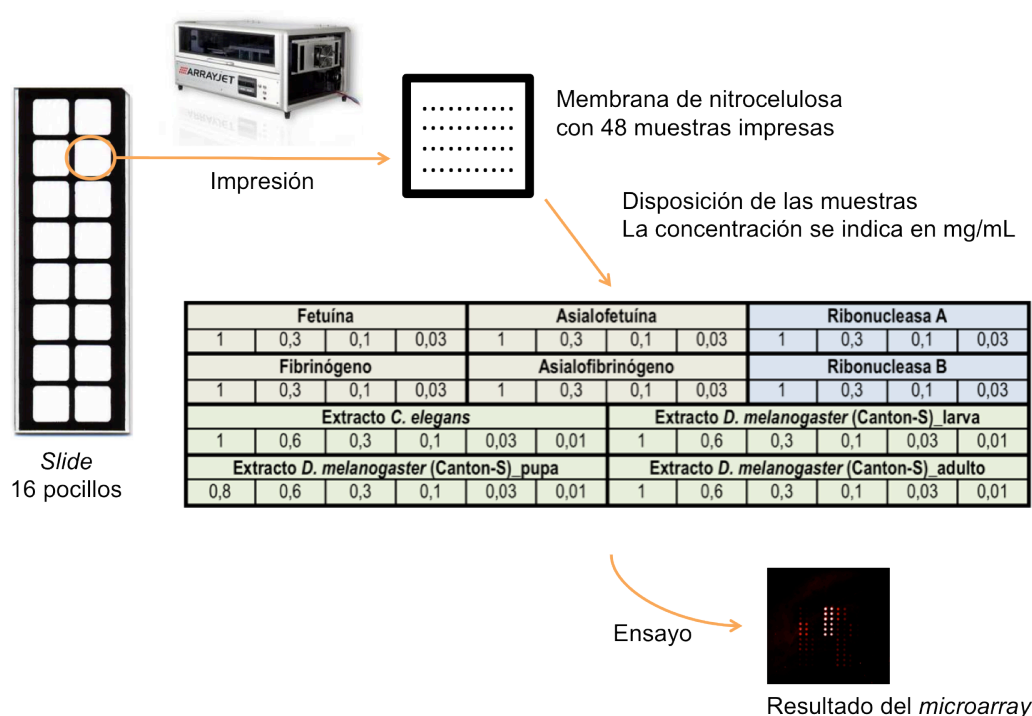
**Figura 3.2.** Reacción de biotinylación donde se muestra la incorporación de la biotina sobre el grupo  $\epsilon$ -amino del aminoácido lisina.

Para comprobar que las lectinas se biotinilaron adecuadamente, se hizo una prueba de *dot blot* en la que se incubaron las proteínas biotiniladas con estreptavidina. En primer lugar, se pincharon 2  $\mu$ L de muestra en una membrana de nitrocelulosa. Después se bloqueó la membrana con TBSB durante 1 hora a temperatura ambiente. Tras varios lavados con TBS, la membrana se incubó con estreptavidina-HRP a 1  $\mu$ g/mL en TBSI durante 35 minutos a temperatura ambiente. En último lugar y tras un lavado con TBS, se reveló la membrana con el kit SIGMAFAST DAB with Metal Enhancer (Sigma-Aldrich), descrito anteriormente (sección 3.6. de *Materiales y Métodos*).

### 5.3. Microarrays

El uso de superficies en las que se inmovilizan diferentes glicoproteínas no solo permite estudiar la capacidad de unión de lectinas, sino que también nos ayuda a analizar su especificidad gracias al uso de competidores o inhibidores de la unión (estudios de inhibición).

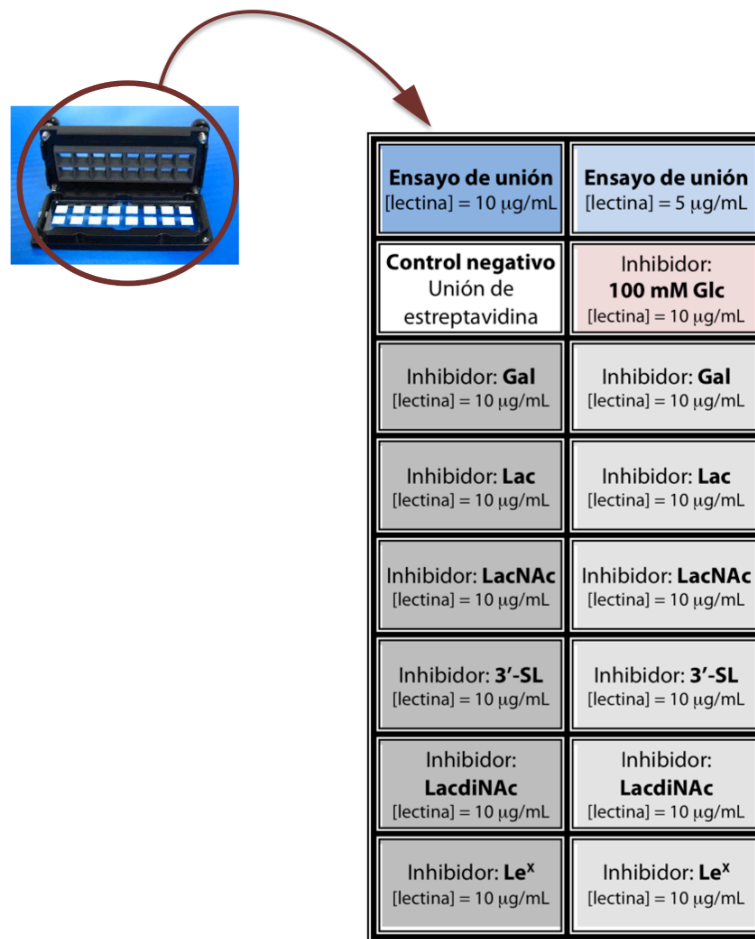
En el presente trabajo se investigó tanto la unión de los módulos lectina LSL<sub>150</sub>, LSL<sub>150</sub>- $\alpha$ F y LBL<sub>152</sub> a algunas glicoproteínas y extractos celulares como la inhibición de esta unión en presencia de determinados azúcares de interés. Para tal fin, se prepararon placas de vidrio o *slides* que contenían 16 pocillos o membranas de nitrocelulosa cada una donde se imprimieron por triplicado 48 condiciones distintas consistentes en una familia de glicoproteínas (fetuína, asialofetuína, fibrinógeno y asialofibrinógeno), extractos celulares procedentes de homogeneizados de *Caenorhabditis elegans* y *Drosophila melanogaster* en diferentes estadios de crecimiento y los respectivos controles negativos (ribonucleasa A y B) (**Figura 3.3**). Las placas de vidrio (Nexterion slide NC-N 16, Schott) se imprimieron usando un Sprint Inkjet Microarrayer (Arrayjet), que trabaja dispensando 100 pL de muestra por *spot* en un rango de 0,03-1 mg/mL (equivalente a 3-100 pg de glicoproteína por *spot*) para las glicoproteínas, y de 0,01-1 mg/mL (1-100 pg/*spot*) para los extractos celulares. Con ayuda de un escáner Gene Pix 4200AL Axon (Molecular Devices) se chequeó la inmovilización, midiendo la señal de emisión del fluoróforo Cy3 (cianina) (Molecular Probes) a una longitud de onda de 532 nm. El Cy3 debe añadirse a las muestras antes de la impresión de las placas.



**Figura 3.3.** Representación esquemática de la preparación de las placas de vidrio (*slides*) donde se indica la distribución de las muestras (por triplicado) de cada una de las membranas de nitrocelulosa del *microarray*. En *marrón claro* están indicadas las glicoproteínas; en *verde*, los extractos celulares; y en *azul*, los controles negativos.

En cada *slide* se llevaron a cabo ensayos tanto de unión como de inhibición. En el caso de los ensayos de unión, las lectinas se prepararon solubles en tampón TSA sin azúcares en el medio. Por el contrario, en los ensayos de inhibición, las lectinas se preincubaron con diferentes azúcares. Para completar el experimento, se incluyó un control de la unión inespecífica de la estreptavidina y un control negativo con glucosa (**Figura 3.4**). Además, dentro de la colección de azúcares que se ensayaron, se empleó lactosa como inhibidor, que al saber previamente que se une a las lectinas, actúa como control positivo. Para llevar a cabo estos ensayos, las placas se bloquearon con TBSB durante 1 hora a temperatura ambiente. Después, fueron lavadas con TBSW (TBS con Tween-20 al 0,05%). Una vez lavadas, las membranas se incubaron con 150  $\mu$ L de lectinas biotiniladas a 10  $\mu$ g/mL durante 90 minutos a temperatura ambiente. Tras nuevos lavados con TBSW, a la membrana se le añadió la estreptavidina a 1  $\mu$ g/mL conjugada con AlexaFluor-647. Esta última etapa transcurrió durante 35 minutos a temperatura ambiente y en oscuridad para no dañar al fluoróforo. Finalmente, las membranas fueron lavadas exhaustivamente con TBSW y agua antes de ser

escaneadas para detectar la señal de emisión del AlexaFluor-647. Las intensidades fueron cuantificadas usando el programa GenePix Pro 6.0 (Molecular Devices).



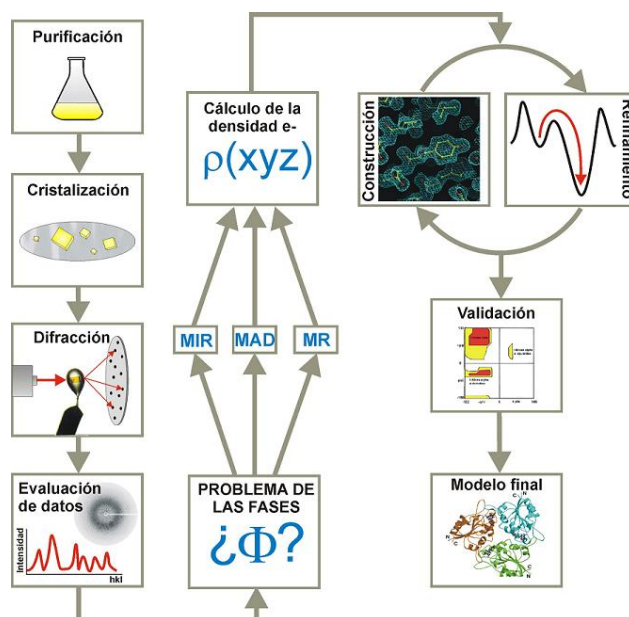
**Figura 3.4.** Esquema en el que se representa cómo se han distribuido los diferentes ensayos en cada placa de vidrio o *slide*. En azul se indican los ensayos de unión de las lectinas a las glicoproteínas y extractos celulares inmovilizados; en blanco, el control de unión de la estreptavidina; en rojo, la inhibición con glucosa, un azúcar no específico de estas lectinas (control negativo); y en gris, los distintos ensayos de inhibición. Los azúcares de la columna de la izquierda (*color gris oscuro*) se prepararon a mayor concentración que los de la derecha (*color gris claro*).

## 6. Estudio estructural de proteínas

El conocimiento del mecanismo de acción biológica de las proteínas implica estudiarlas funcional y estructuralmente. Y es que la caracterización de los detalles específicos de las interacciones que establecen con los correspondientes ligandos, que en última instancia son los que nos van a permitir sacar conclusiones acerca de su mecanismo de actuación en la naturaleza, pasa necesariamente por el conocimiento de sus estructuras tridimensionales (Weis y Drickamer, 1996). Técnicas como la cristalografía de rayos X aportan informa-



ción valiosa tanto de la especificidad de las proteínas como de las relaciones entre las diferentes familias estructurales de lectinas (Loris, 2002).



**Figura 3.5.** Esquema general de las distintas etapas que configuran un estudio estructural mediante cristalografía de rayos X.

## 6.1. Cristalografía de proteínas

La Biología Estructural se vale de diferentes técnicas que permiten caracterizar estructuralmente a las proteínas. Las tres más importantes son la cristalografía de rayos X, la resonancia magnética nuclear (RMN) y la microscopía electrónica. Sin embargo, la cristalografía es la más exitosa si nos atenemos al número de estructuras depositadas actualmente en el *Protein Data Bank* (PDB) (<http://www.rcsb.org/pdb/home/home.do>).

La cristalografía permite resolver la estructura de proteínas independientemente de su masa molecular, lo que supone una importante ventaja frente a la RMN, y alcanzar resoluciones atómicas, a diferencia de lo que ocurre con la microscopía electrónica. No obstante, uno de los principales escollos de la cristalografía es el requisito de obtener cristales de proteína de alta calidad (Gómez-Moreno y Sancho, 2003).

La resolución estructural de proteínas mediante cristalografía de rayos X implica una serie de pasos experimentales hasta la obtención de un modelo estructural final (**Figura 3.5**). El análisis del modelo, así como la caracterización funcional de la proteína, permiten establecer la relación estructura-función, el objetivo principal de la Biología Estructural.

### 6.1.1. Cristalización

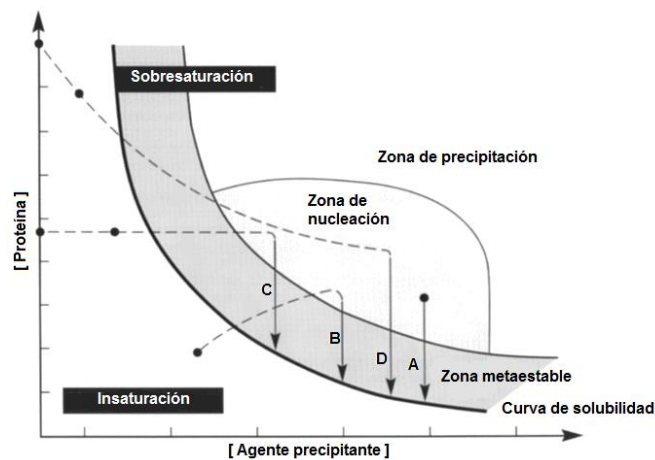
Como muchas otras moléculas, las proteínas pueden formar cristales como resultado de su ordenación regular en el espacio mediante el establecimiento de interacciones débiles no covalentes entre ellas (McPherson, 2004). Actualmente, aunque entendemos globalmente el proceso de cristalización desde una perspectiva fisicoquímica como una transición de fase entre una fase soluble y una fase sólida cristalina, no podemos predecir si una proteína es cristalizable, ni las condiciones en las que lo hará, en caso de hacerlo, tomando como punto de partida información derivada de la propia proteína. No obstante, de un modo general, podemos afirmar que para que una proteína cristalice, su grado de pureza debe ser muy alto ( $> 95\%$ ), además de ser homogénea química y conformacionalmente. No es extraño, por tanto, que las etapas de producción y purificación sean críticas para la calidad final de los cristales.

Los cristales de proteína se caracterizan por tener un alto contenido de solvente acuoso desordenado (entre un 30 y un 80%) localizado en los canales de dichos cristales (McPherson, 2004). Este desorden intrínseco debido al solvente es el responsable de que los cristales de macromoléculas no suelen difractar a resolución atómica ( $> 1,2 \text{ \AA}$ ), pero, por otro lado, este alto contenido de solvente hace que las proteínas se encuentren en el cristal en un medio esencialmente acuoso que preserva su estado nativo. De hecho, cuando ha sido posible, la adición de sustratos sobre cristales de enzimas ha rendido los productos esperados, lo que demuestra que la proteína es activa.

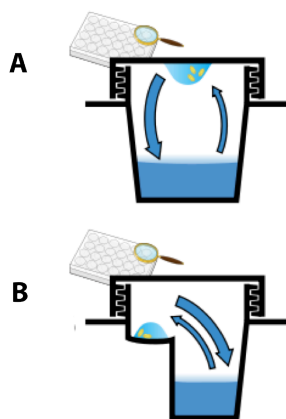
La obtención de cristales se consigue a partir de disoluciones sobresaturadas de proteína en presencia de un agente precipitante. Las diferentes aproximaciones experimentales en cristalografía tienen como objetivo conducir a la solución de proteína a la denominada región de nucleación (**Figura 3.6**). En esta región del diagrama de fases, se produce la formación de núcleos cristalinos (nucleación), los cuales sirven de puntos de partida para el posterior crecimiento de los cristales. La formación de cristales, por lo tanto, ocurre en dos etapas diferenciadas: la nucleación y el crecimiento.

En los casos positivos de cristalización, el sistema proteína-solución precipitante viajaría por el diagrama de fases desde un estado de sobresaturación, pasaría por la región de nucleación, donde se generarían unos pocos núcleos, llegando hasta la zona metaestable en donde se produciría crecimiento cristalino, es decir, la incorporación de moléculas de proteína a los núcleos desde la fase soluble, de forma ordenada. Los agregados amorfos y

otros precipitados no cristalinos ocurren en la denominada región de precipitación (**Figura 3.6**).



**Figura 3.6.** Diagrama de fases o diagrama de Ostwald en el que se muestran las diferentes zonas en las que la proteína se puede encontrar durante un experimento de cristalización. Además, se muestran las rutas que sigue la muestra en función del método de cristalización usado: cristalización en baño o *batch* (A), difusión de vapor (B), diálisis (C) y contradifusión en fase líquida (D).



**Figura 3.7.** Representación esquemática del dispositivo experimental empleado para la cristalización mediante difusión de vapor por gota colgante (A) y gota sentada (B). La gota que se encuentra lejos del reservorio contiene solución de cristalización y proteína. En los casos favorables, después de alcanzarse el equilibrio, los cristales de proteína aparecerán en su interior.

Aunque existen muchos métodos para tratar de obtener cristales, la difusión de vapor es la técnica más común (**Figura 3.7**). En este caso, la solución que contiene la proteína y el agente precipitante se equilibra en un recinto estanco con un reservorio acuoso donde solamente hay solución de cristalización. Normalmente, los volúmenes empleados son 1  $\mu\text{L}$  de solución de proteína mezclado con 1  $\mu\text{L}$  de solución precipitante y 500  $\mu\text{L}$  de solución precipitante en el reservorio. La mezcla con la proteína puede encontrarse suspendida en una gota adherida a un cubreobjetos siliconizado (difusión de vapor en gota colgante) o puede encontrarse sobre una elevación plana o cóncava por encima del nivel de la solución de cristalización del reservorio (difusión de vapor en gota sentada). En ambos casos, se produce una desecación de la gota de proteína por la difusión de vapor de agua y otros

componentes volátiles desde la gota en donde se encuentra la proteína a la solución del reservorio, cuya concentración de precipitante es mayor.

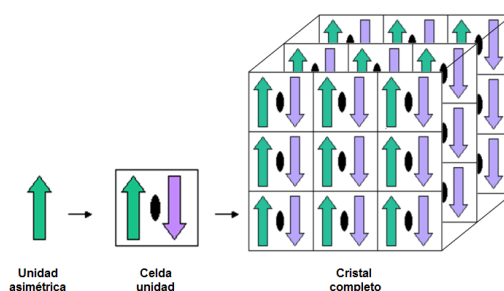
Los factores más comunes que afectan a la solubilidad de las proteínas son el pH, la temperatura, la presencia o no de sales, la adición de solventes orgánicos, el uso de precipitantes poliméricos y la adición de ligandos, entre otros. Debido al alto número de variables de las que depende la formación de cristales, la cristalización de una proteína se afronta como un experimento de prueba y error. La concentración de sales es uno de los parámetros más importantes que deben ajustarse. Cuando la concentración de sal es muy elevada puede darse el fenómeno conocido como *salting out*. Los iones secuestran moléculas de agua, incluyendo aguas estructurales, propiciando la interacción hidrofóbica de las proteínas. Por el contrario, el fenómeno opuesto, *salting in*, hace referencia al aumento de la solubilidad de las proteínas al aumentar la concentración de sal, cuando la concentración de esta última es relativamente baja.

La obtención de cristales de proteína en presencia de ligandos o sustratos se utiliza para conocer los detalles de la interacción proteína-ligando, estudiar el mecanismo de acción o validar ciertas moléculas como potenciales fármacos, entre otras aplicaciones. Para la cristalización de complejos proteína:ligando existen dos estrategias principalmente (Hassell *et al.*, 2007). En la inmersión o *soaking*, los cristales de proteína se sumergen en una solución que contiene el ligando a alta concentración. Como los cristales de proteína presentan grandes canales de solvente, la difusión del ligando hasta el sitio de unión no suele ser un factor limitante. El inconveniente de esta técnica es que en muchas ocasiones este acceso no es posible esencialmente debido al empaquetamiento cristalográfico, o bien porque su unión a la proteína provoca cambios estructurales en la misma que suelen ser incompatibles con la ordenación del cristal. En la cocrystalización, sin embargo, el ligando se añade en la gota al comienzo del experimento de cristalización, o bien se incubaba previamente con la proteína. En general, esta modalidad se realiza utilizando las condiciones de cristalización de la proteína sola, si bien no es extraño comprobar que los complejos cristalizan en condiciones diferentes a las de la proteína libre.

### 6.1.2. Difracción de rayos X

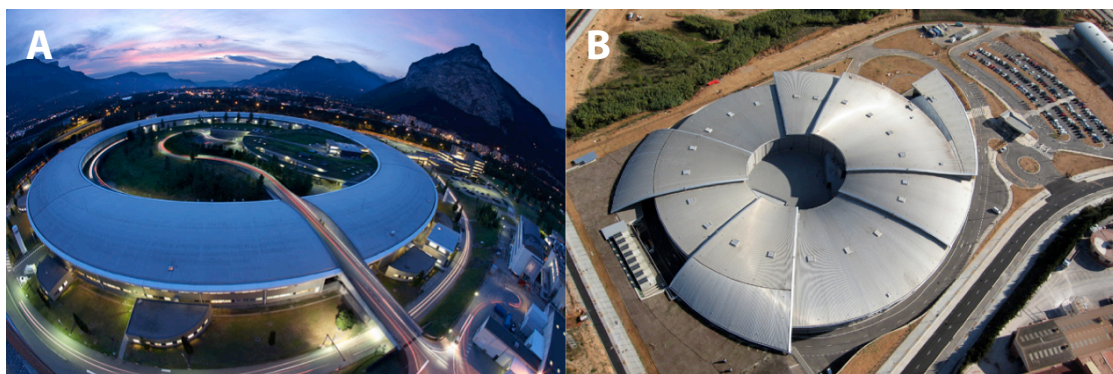
Conseguir que la proteína cristalice es condición *sine qua non* para abordar la resolución estructural de una proteína mediante difracción de rayos X. Un cristal es el resultado de la repetición en tres dimensiones de una unidad básica, conocida como celda unidad. Cada celda unidad se caracteriza por tres ejes ( $a$ ,  $b$  y  $c$ ) y tres ángulos cristalográficos ( $\alpha$ ,  $\beta$

y  $\gamma$ ). A su vez, en la celda unidad definimos la denominada unidad asimétrica (**Figura 3.8**). La aplicación de las operaciones de simetría que definen el grupo espacial del cristal a dicha unidad asimétrica, genera la celda unidad completa. El objetivo de la cristalografía de proteínas es, por tanto, la determinación de la estructura de la unidad asimétrica del cristal analizado.



**Figura 3.8.** La unidad asimétrica se repite dentro de la celda unidad. La repetición de la celda unidad genera el cristal.

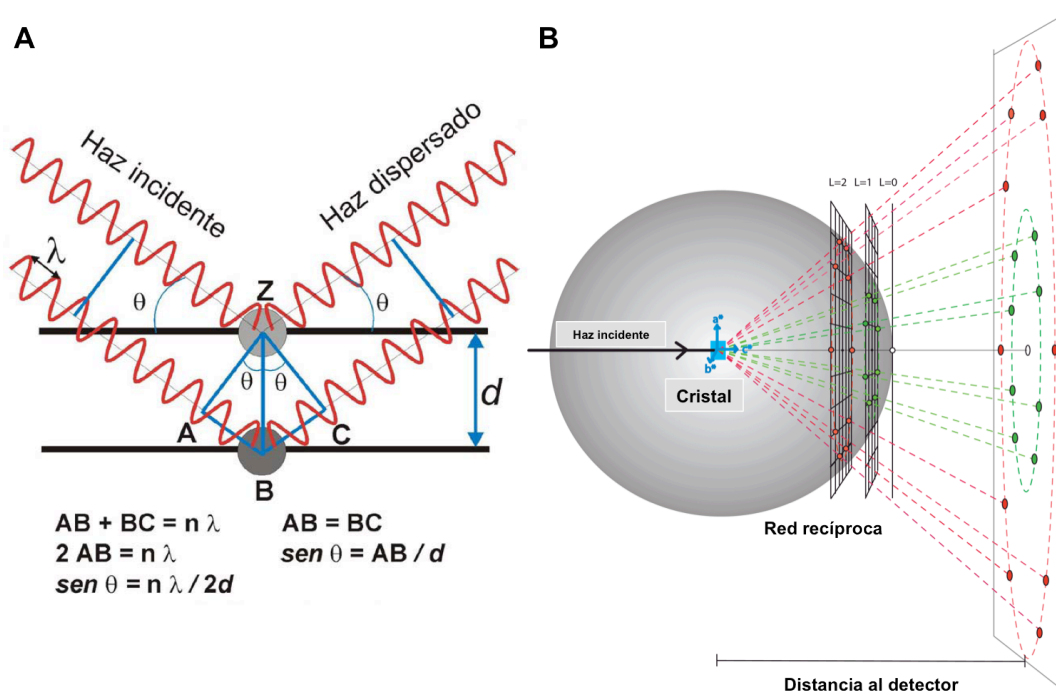
El otro elemento importante para completar un experimento de cristalografía son los rayos X. Los rayos X son fotones con longitudes de onda ( $\lambda$ ) del orden de los ångström (Å), comparables a las distancias interatómicas. La radiación X se puede obtener a partir de generadores de ánodo rotatorio, que tienen un ánodo metálico (cobre, por ejemplo) sobre el que chocan electrones acelerados procedentes de un filamento metálico. También se pueden producir rayos X en las grandes instalaciones de tipo sincrotrón, donde se aceleran electrones en grandes anillos a velocidades cercanas a las de la luz (**Figura 3.9**). Los cambios de dirección que sufren los electrones a lo largo de su trayectoria circular se traducen en la emisión de radiación sincrotrón<sup>33</sup>.



**Figura 3.9.** En **A** se muestra una fotografía del *European Synchrotron Radiation Facility* (ESRF) (Grenoble, Francia); mientras que en **B**, se muestra una fotografía del *ALBA Synchrotron Light Facility* (Sardañola del Vallés, España). Ambos son pequeños aceleradores de partículas donde se genera radiación sincrotrón.

Cuando los rayos X inciden sobre un cristal de proteína se produce el fenómeno conocido como difracción de rayos X, fundamento físico de la cristalografía. Los rayos X interactúan con la materia cristalina, en particular con los electrones, mediante distintos mecanismos. Solo uno de ellos, la interacción elástica o dispersión de Thomson, es el responsable en última instancia del fenómeno de la difracción. La dispersión Thomson descansa sobre la base de que el fenómeno subyacente es una interacción partícula (electrón) onda (rayos X), tratada clásicamente, sin invocar a conceptos mecano cuánticos.

Por otro lado, la aproximación fenomenológica de la difracción de rayos X por un cristal considera este fenómeno como resultado de procesos de reflexión de haces de rayos X con una longitud de onda determinada ( $\lambda$ ) que inciden con un ángulo de incidencia dado ( $\theta$ ) sobre planos del cristal (planos  $hkl$ ) separados una determinada distancia entre sí (espacio interplanar,  $d$ ). Es en este contexto donde reside la denominada ley de Bragg (**Ecuación 3.10**), que impone las condiciones necesarias para que se produzca un fenómeno de interferencia constructiva por los rayos X reflejados por una determinada familia de planos  $hkl$  (**Figura 3.10A**) (Rhodes, 2006).



**Figura 3.10.** Interpretación geométrica de la difracción de rayos X. **A.** Condiciones que se cumplen en la difracción de rayos X sobre un cristal ordenado. Los dos haces incidentes que penetran en el cristal son reflejados de forma coherente, interfiriendo constructivamente si se satisface la ley de Bragg. **B.** Representación de la esfera de Ewald. Cuando un punto recíproco se encuentra en la superficie de la esfera de Ewald se cumple la ley de Bragg, es decir, se satisface la condición necesaria de interferencia constructiva (líneas rojas y verdes).

Una de las interpretaciones geométricas de la ley de Bragg fue originariamente propuesta por Ewald, la cual hace uso de la denominada esfera de Ewald (**Figura 3.10B**). Este formalismo permite comprender de un modo sencillo la ley de Bragg en tres dimensiones. Sin entrar en detalles, la condición de interferencia constructiva, es decir, el cumplimiento de la ley de Bragg, ocurre cuando un punto de la red recíproca ( $hkl$ ) se encuentra en la superficie de la esfera de Ewald. Esta representación, a su vez, permite comprender intuitivamente el método de rotación para la toma de datos de difracción (ver apartado 6.1.3 de *Materiales y Métodos*).

$$2d_{hkl}\sin\theta = n\lambda \quad \text{(Ecuación 3.10)}$$

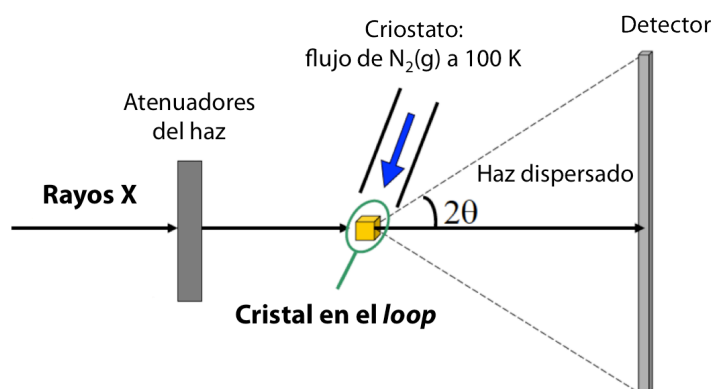
Debido a la fragilidad de los cristales de proteína, estos sufren daño cuando son expuestos de forma prolongada a una fuente de rayos X (Garman, 1999; Garman y Owen, 2006). Se genera calor y se forman radicales libres que dañan las moléculas irreversiblemente. Por esta razón, los cristales se mantienen bajo un flujo de nitrógeno a una temperatura próxima a los 100 K. La utilización de estas temperaturas permite incrementar la vida media del cristal y la resolución de los datos de difracción, ya que reduce la vibración térmica y puede fijar conformaciones en zonas flexibles de la proteína (Garman, 1999). Antes de someterlo a baja temperatura, el cristal ha de sumergirse en una solución crioprotectora (Rubinson *et al.*, 2000) que evita la formación de cristales de hielo que puedan deteriorar los cristales de proteína o interferir con el patrón de difracción del cristal de proteína. Los cristales se cogen directamente de la gota donde han crecido con ayuda de un lazo de nylon (*loop*), el cual es transparente a los rayos X. El lazo con el cristal es rápidamente sumergido en una gota con solución crioprotectora donde se mantiene durante cierto tiempo, normalmente, unos pocos segundos. En algunas ocasiones, los cristales aparecen directamente en condiciones de cristalización que son crioprotectoras de por sí debido a la presencia de algún agente como glicerol, PEG 400 o algunas sales, entre otros. En este caso, el cristal se puede congelar directamente después de ser recogido.

### 6.1.3. Obtención de los datos de difracción y su análisis preliminar

Los patrones de difracción que se obtienen cuando se hace incidir un haz de rayos X sobre un cristal de proteína van a ser los datos experimentales para iniciar el proceso de resolución estructural (**Figura 3.11**). La toma de datos se ha llevado a cabo mediante el método de rotación (ver apartado 6.1.2 de *Materiales y Métodos*), lo cual implica desarrollar una estrategia de rotación del cristal respecto del haz incidente de rayos X que haga



pasar a través de la superficie de la esfera de Ewald a los puntos de la red recíproca que definen la unidad asimétrica en el espacio recíproco (Figura 3.10B). La etapa de recogida de datos es muy importante, pues la calidad de estos va a condicionar todo el proceso posterior de resolución estructural. La recogida se inicia normalmente con la toma de una o varias imágenes separadas un cierto ángulo (típicamente dos imágenes separadas  $90^\circ$ ), utilizando  $1^\circ$  de rotación. La inspección de estas imágenes permite diagnosticar numerosos aspectos del cristal y su difracción: resolución a la que el cristal puede difractar, la anisotropía, la presencia de maclas o de varios cristales, anillos de hielo, posible solapamiento y forma de las reflexiones etc. Este proceso inicial y su análisis con determinadas herramientas (Popov y Bourenkov, 2003; Bourenkov y Popov, 2006) nos permitirá diseñar una estrategia de recogida de datos, incluyendo los tiempos de exposición, el ángulo de giro por imagen ( $\Delta\varphi$ ) e intervalo angular total, la distancia cristal-detector, etc. (Evans, 1999).



**Figura 3.11.** Esquema de un experimento de recogida de datos de difracción de rayos X. El cristal difracta la radiación incidente en muchos haces discretos, cada uno de los cuales produce un *spot* o reflexión en el detector. La posición y la intensidad de estas reflexiones contienen parte de la información necesaria para la determinación posterior de las estructuras moleculares.

El procesamiento de los datos de difracción consta de cuatro etapas: (1) autoindexación, (2) integración, (3) escalado y (4) cálculo de los factores de estructura. En primer lugar, durante la autoindexación se determinan los parámetros de la celda unidad, la orientación del cristal y su mosaicidad, así como proporcionar una hipótesis inicial sobre el grupo puntual y grupo espacial. En segundo lugar, la integración nos permite estimar las intensidades de las reflexiones acompañadas de su correspondiente error o desviación estándar relativa. Como resultado, se obtiene un conjunto de datos con todas las reflexiones medidas ( $hkl$ ), sus intensidades ( $I_{hkl}$ ) y sus errores asociados ( $\sigma_{hkl}$ ). Posteriormente, en la etapa de escalado se normalizan las intensidades, minimizando las diferencias experimentales entre los valores de intensidades correspondientes a una misma reflexión, o equivalentes por



simetría, medidas en diferentes imágenes. Además, se obtienen una serie de parámetros estadísticos que permiten evaluar la calidad de los datos de difracción en función de la resolución. Entre estos, destacan la multiplicidad (representa el número de veces que se ha medido una reflexión), la completitud (porcentaje del espacio recíproco que se ha recogido durante el experimento), la relación señal-ruido ( $I/\sigma$ ) e indicadores de la calidad de los datos ( $R_{merge}$ ,  $R_{meas}$  y  $R_{pim}$ ) (Evans, 2011) (**Figura 3.12**). Finalmente, una vez determinadas las intensidades de las reflexiones, se calculan los módulos de los factores de estructura, cuyo cuadrado es proporcional a la intensidad.

$$R_{merge} = \frac{\sum_{hkl} \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^n I_i(hkl)} \quad (\text{Ecuación 3.11})$$

$$R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^n I_i(hkl)} \quad (\text{Ecuación 3.12})$$

$$R_{pim} = \frac{\sum_{hkl} \sqrt{\frac{1}{n-1}} \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^n I_i(hkl)} \quad (\text{Ecuación 3.13})$$

**Figura 3.12.** Expresiones matemáticas que describen a los indicadores usados para evaluar la calidad de los datos durante la fase de escalado.  $I_i(hkl)$  hace referencia a la intensidad de una reflexión puntual, mientras que  $\bar{I}(hkl)$  se corresponde con la intensidad media. Cabe destacar que tanto el  $R_{meas}$  como el  $R_{pim}$  no se ven influenciados por la multiplicidad, ya que  $n$  (número de veces que se observa una reflexión determinada) pondera este efecto.

Un aspecto importante en el procesamiento de datos cristalográficos es elegir en qué punto cortar la resolución para un set de datos determinado. Tradicionalmente, los indicadores de calidad de la Figura 3.12 han sido utilizados con este propósito, ajustándose la resolución a la concha cuyo valor de  $R_{merge}$  (o  $R_{meas}$ ) no excediera más de ~0,6-0,8 y cuya relación señal-ruido ( $\langle I/\sigma(I) \rangle$ ) se mantuviera alrededor de ~2. Sin embargo, esta práctica tan extendida ha sido recientemente puesta en entredicho, demostrándose recientemente que la utilización del coeficiente de correlación de Pearson ( $CC$ ) es una herramienta mucho más adecuada para este fin (Karplus y Diederichs, 2012). Concretamente, se definen los coeficientes  $CC_{1/2}$  y  $CC^*$  (**Ecuación 3.14**) que asumen que si dividimos un set de datos en dos partes, los errores en ambas son al azar y similares en tamaño.  $CC_{1/2}$  presenta valores de 1,0 a baja resolución y se aproxima a 0,1 a resoluciones más altas. Además, estos coeficientes pueden compararse directamente con los respectivos coeficientes  $CC_{work}$  y  $CC_{free}$ , determinados en la etapa de refinamiento y reconstrucción del modelo final. Éstos nuevos indicadores ya están siendo incorporados en las últimas actualizaciones de los programas de procesamiento cristalográfico.

$$CC^* = \sqrt{\frac{2CC_{1/2}}{1+CC_{1/2}}} \quad \text{(Ecuación 3.14)}$$

#### 6.1.4. Resolución estructural y validación

En un experimento de cristalografía de rayos X, el objetivo que se persigue es la determinación de la densidad electrónica  $\rho(xyz)$  (Ecuación 3.15) en cada punto  $xyz$  de la unidad asimétrica. Como se puede ver en esta ecuación, para ello debemos conocer tanto la amplitud ( $|F_{hkl}|$ ) como la fase  $\alpha(hkl)$  de cada factor de estructura. El denominado problema de la fase en cristalografía reside precisamente en que las medidas de difracción no aportan información sobre las fases, sólo permiten estimar la amplitud de los factores de estructura. En el caso de la determinación estructural de proteínas, son tres los métodos más utilizados para resolver el problema de la fase: reemplazo isomorfo simple/múltiple (SIR/MIR), técnicas basadas en medidas de señal anómala a una o múltiples longitudes de onda (SAD/MAD) y reemplazo molecular<sup>34</sup> (MR). La relación entre fases  $[\alpha(hkl)]$  y las amplitudes  $[|F_{hkl}|]$  de los factores de estructura con la densidad electrónica en un punto  $xyz$  de la celdilla unidad viene dada por esta ecuación, cuya representación proporciona lo que se denomina mapa de densidad electrónica.

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F_{hkl}| e^{-2\pi i(hx+ky+lz-\alpha(hkl))} \quad \text{(Ecuación 3.15)}$$

Una vez resuelto el problema de la fase, el modelo inicial se somete a varios ciclos de reconstrucción manual trabajando *in situ* sobre el mapa de densidad y de refinamiento, lo que implica el ajuste de las coordenadas ( $xyz$ ) de cada átomo y del factor térmico  $B^{13}$ . Durante el refinamiento, se imponen restricciones estereoquímicas, las cuales permiten mantener distancias y ángulos de enlace correctos. El indicador estándar que permite evaluar el acuerdo entre el modelo refinado con los datos es el factor  $R$  (Ecuación 3.16).  $R$  es cero para un acuerdo perfecto con los datos y alrededor de 0,59 para un modelo aleatorio. Puesto que el valor de  $R$  puede hacerse arbitrariamente pequeño con modelos con un número lo suficientemente grande como para sobrefinarse los datos, Brünger introdujo el parámetro  $R_{free}$ , un factor  $R$  de validación cruzada que se calcula con un 5-10% de las reflexiones tota-

les, que no son incluidas en el refinamiento. El parámetro  $R_{work}$  se calcula con el resto de las reflexiones.

$$R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \quad \text{(Ecuación 3.16)}$$

La etapa de validación del modelo constituye la última etapa de la resolución estructural. Se comprueba tanto la estereoquímica del modelo como la correlación de la densidad electrónica con el modelo atómico. La validación de la estereoquímica implica la comprobación de las distancias y ángulos de enlace (tanto de la cadena principal como de las laterales), las conformaciones de los rotámeros y los ángulos de torsión de la cadena principal ( $\varphi$  y  $\psi$ ). La representación de los aminoácidos en el diagrama de Ramachandran permite identificar los residuos que se encuentran en zonas no permitidas, siendo necesario que el número de residuos fuera de las zonas permitidas sea inferior al 2% (Ramachandran *et al.*, 1963).

Una vez que la construcción del modelo se ha completado, el siguiente paso consiste en validar y depositar las coordenadas en el *Protein Data Bank* (PDB). Durante esta etapa, el PDB cuenta con una herramienta de validación, llamado ADIT (AutoDep Input Tool), que revisa la consistencia del modelo atómico, su geometría y las interacciones comparando el modelo con otras estructuras estándar ya conocidas. Los contactos entre moléculas de la unidad asimétrica o relacionadas por simetría, la longitud de enlace, los ángulos de enlace y torsión, entre otros, son chequeados en detalle. Aguas posicionadas incorrectamente, errores de quiralidad y discordancias con la secuencia son elementos que también se resaltan. ADIT se vale de los programas NUCHECK, PROCHECK (Laskowski *et al.*, 1993) y SFCHECK (Vaguine *et al.*, 1999) para la validación exhaustiva de las coordenadas durante su deposición. Esta validación rigurosa actúa como un elemento externo de seguridad que confirma que las estructuras depositadas cumplen todos los requerimientos exigidos.

### 6.1.5. Caracterización estructural de las proteínas de esta tesis

La primera aproximación para llevar a cabo la cristalización de las proteínas fue ajustar su concentración, pues ésta es una variable crítica en cualquier experimento de cristalización. El empleo del kit PCT (Pre-crystallization Test, Hampton Research) permitió comprobar qué concentración podríamos usar de partida y así reducir desde un principio la cantidad de proteína de cara a la posterior optimización de las condiciones de cristaliza-

ción. La apariencia de la gotas (gotas claras, aparición de precipitados, etc.) nos revela en qué rango de concentraciones nos debemos mover.

Una vez seleccionada la concentración de proteína, se utilizó un robot Innovadyne Nanodrop ExtY (Solve Scientific), que dispensa gotas automáticamente sobre placas Innovaplate SD-2 de 96 pocillos (Solve Scientific) para llevar a cabo múltiples ensayos de cristalización mediante difusión de vapor en gota sentada. El volumen final de las gotas se ajustó a 500 nL, resultado de mezclar 250 nL de solución de proteína con 250 nL de solución de cristalización, y se equilibró frente a 65  $\mu$ L de solución de cristalización presente en cada uno de los pocillos. De esta manera, se pueden probar rápidamente cientos de soluciones diferentes con el fin de explorar gran parte del espacio de cristalización. Las soluciones de cristalización que fueron empleadas en esta etapa fueron soluciones incluidas en los siguientes kits comerciales: JBS Classic 1-4 (Jena Bioscience), Crystal Screen 1 y 2 (Hampton Research), Index (Hampton Research), SaltRx (Hampton Research), PACT (Qiagen) y JCSG (Qiagen).

El proceso de optimización comenzó con el escalado de las gotas: el experimento de cristalización en microplaca se trasladó a una placa Linbro de 24 pocillos (Hampton Research) en donde el volumen del reservorio se aumentó a 500  $\mu$ L a la vez que se probaron diferentes relaciones de volumen de proteína:precipitante. Posteriormente, se procedió a preparar las soluciones de cristalización manualmente para así comprobar cómo responde el sistema cuando se varían finamente variables como la concentración de precipitante, sal y proteína, el valor del pH y/o la temperatura. Además, en algunos casos se añadieron aditivos, tales como detergentes, metales o ligandos que pueden llegar a mejorar la calidad final de los cristales. En todos los ensayos del escalado se utilizó para cristalizar el método de difusión de vapor en gota colgante.

Para corroborar que los cristales obtenidos tanto en las etapas iniciales como en el posterior escalado son de proteína y no de ningún otro compuesto presente en la solución, se usó el compuesto Izit Crystal Dye (Hampton Research) que tiñe selectivamente los cristales de proteína gracias a que puede penetrar por los grandes canales acuosos presentes en su interior.

Una vez que los cristales se hubieron optimizado, se seleccionó un crioprotector adecuado, a menos que la solución de cristalización ya contuviera algún agente con capacidad crioprotectora de por sí. La elección del agente crioprotector se basó en criterios experimentales, considerando que el mejor va a ser aquél que permita obtener un patrón de difracción de mayor calidad. Por lo general, se utilizó glicerol y 2-metil-2,4-pentanodiol

(MPD) a concentraciones próximas al 20-25% (v/v). Antes de congelar los cristales, éstos fueron transportados a una solución que contenía los mismos componentes que la solución de cristalización más el agente crioprotector. En esta solución se mantuvieron durante no más de 10 segundos hasta que finalmente se sumergieron en nitrógeno líquido.

La toma de datos de difracción se llevó a cabo en estaciones de tipo sincrotrón, tanto en el ESRF (*European Synchrotron Radiation Facility*, Grenoble, Francia) como en ALBA (*ALBA Synchrotron Light Facility*, Sardañola del Vallés, España). Los cristales fueron medidos en diferentes líneas destinadas a macromoléculas biológicas en las que existe un *software* específico que permite programar recogidas de datos en función de la calidad de la difracción. Este *software* está constituido principalmente por los siguientes programas: mxCuBE (permite centrar el haz sobre el cristal y programar la recogida), EDNA (permite hacer un análisis preliminar de los datos y calcular una estrategia de recogida) y ADXV (visualizador de los patrones de difracción).

El procesamiento de datos se hizo con ayuda de diversos programas de análisis cristalográfico. Inicialmente, la indexación y la integración de las intensidades obtenidas directamente de los patrones de difracción se llevó a cabo con el programa iMosflm (Leslie, 2006) o XDS (Kabsch, 2010). Los procesos de escalado y posterior cálculo de los factores de estructura se hicieron usando SCALA (Evans, 2006), programa que se encuentra dentro del paquete de software CCP4 (Cowtan *et al.*, 2011; Winn *et al.*, 2011). De forma general, se utilizó el programa POINTLESS (Evans, 2006) para corroborar que el grupo espacial escogido inicialmente era el correcto.

Las estructuras obtenidas en este trabajo se resolvieron aplicando el método de reemplazo molecular ya que existía como modelo de reemplazo la estructura de la lectina LSLa, que había sido resuelta recientemente (Mancheño *et al.*, 2005). Para tal fin, se utilizaron distintos programas, que incluyeron MOLREP (Vagin y Teplyakov, 2010) y PHASER (McCoy *et al.*, 2007). Una vez obtenida la solución del reemplazo, los ciclos de refinamiento y reconstrucción se llevaron a cabo usando PHENIX (Adams *et al.*, 2010) y Coot (Emsley y Cowtan, 2004), respectivamente. Finalmente, la validación se realizó con el programa MOLPROBITY (Davis *et al.*, 2007).

## 6.2. Métodos de predicción de estructura terciaria

Cuando experimentalmente no es posible obtener la estructura tridimensional de una proteína, existen aproximaciones bioinformáticas que permiten predecir con mayor o menor confianza su estructura. En términos generales, la estrategia se basa en el análisis de las

relaciones entre las estructuras y secuencias de proteínas homólogas con la secuencia de la proteína diana. El modelado por homología funciona cuando el porcentaje de identidad de secuencia es relativamente alto. Generalmente, si las proteínas comparten aproximadamente el 50% de la secuencia, el modelo no suele presentar un RMSD de los C $\alpha$  superior a  $\sim 1$  Å (Bordoli *et al.*, 2009).

Para modelar las estructuras de las lectinas SSL y ANL, se utilizaron los programas SWISS-MODEL (Bordoli *et al.*, 2009) y Robetta ([www.robetta.bakerlab.org](http://www.robetta.bakerlab.org)) (Kim *et al.*, 2004). El primero de ellos funciona con alineamientos de secuencia con proteínas homólogas con altos porcentajes de identidad de secuencia. En nuestro caso, se hicieron alineamientos entre las lectinas y el módulo LSL<sub>150</sub> con ayuda del programa ClustalW2 (Larkin *et al.*, 2007) del servidor ExPASy. El segundo de los programas de modelado, sin embargo, realiza una búsqueda contemplando el uso de homólogos lejanos.

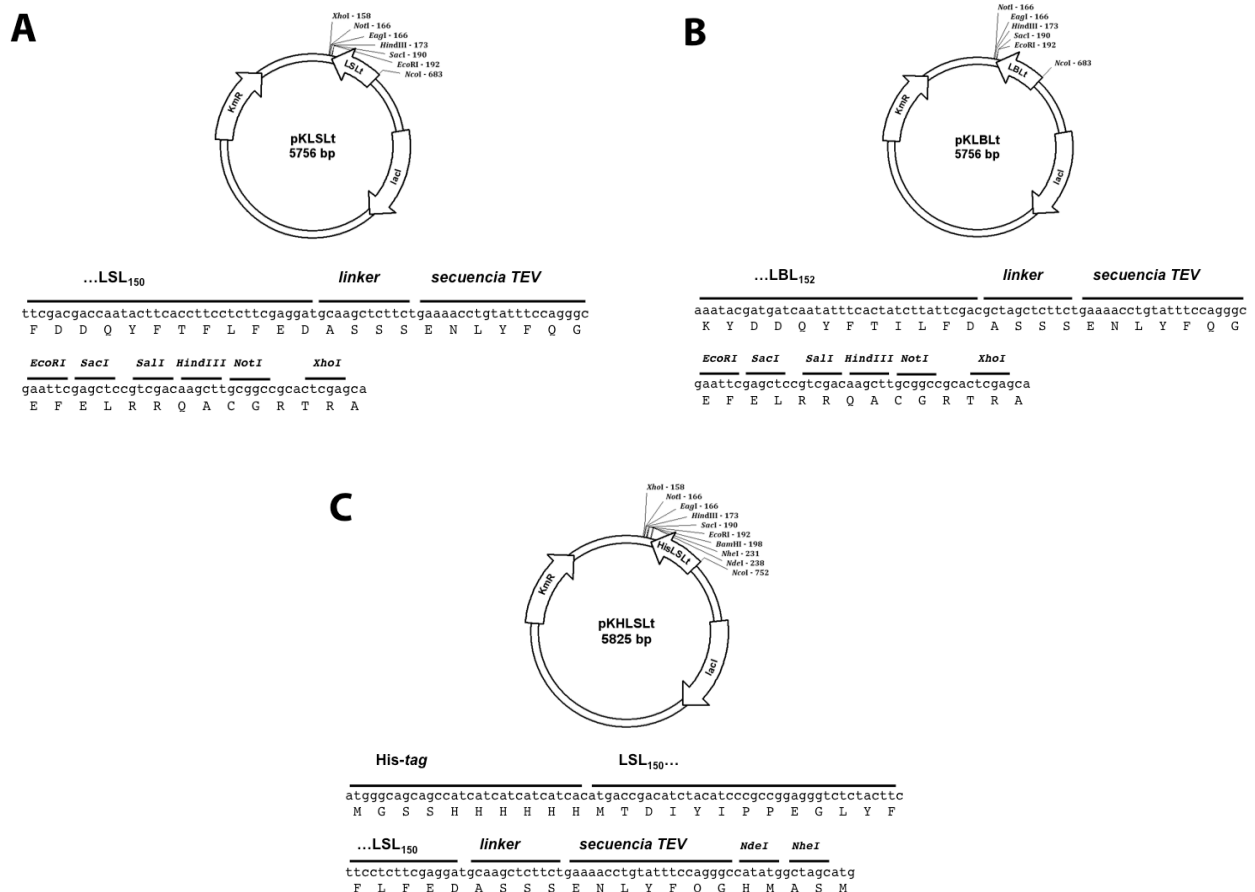
## 7. Aplicaciones biotecnológicas de las lectinas con estructura tipo trébol $\beta$

### 7.1. Producción de proteínas recombinantes conteniendo etiquetas de fusión basadas en módulos lectina N-terminales

La construcción de los vectores pKLSLt, pKLBLt, pKELSLt y pKHLSLt (**Figura 3.13**) se llevó a cabo usando como molde el vector comercial pET-28a(+) (Novagen), por lo que todos ellos presentan el promotor de la polimerasa del fago T7, resistencia al antibiótico kanamicina (Km<sup>R</sup>) y la capacidad de expresar el represor del operón lactosa (*lacI*), que se inhibe bajo la presencia de IPTG. El método seguido para la construcción de los vectores y los pares de oligonucleótidos utilizados quedan recogidos en la **Tabla 3.4**. Cabe destacar que los cuatro vectores presentan en sus extremos 3' una secuencia que codifica un segmento flexible, conocido como secuencia *linker* (aminoácidos ASSS), y un sitio de reconocimiento por la endoproteasa del virus del grabado del tabaco (TEV), cuya secuencia de reconocimiento consenso en aminoácidos es ENLYFQG. De esta manera, una vez producida la proteína de fusión, la lectina puede ser eliminada mediante una digestión controlada con la proteasa TEV (Angulo *et al.*, 2011).

Los vectores de expresión fueron diseñados de tal manera que se pudiesen clonar en ellos los genes mediante dos métodos distintos: uno dependiente de enzimas de restricción y ligasa (conserva la región *polylinker* (sitio de multiclonación) en el extremo 3' de la se-

cuencia de reconocimiento de la proteasa TEV, **Figura 3.13**) y otro independiente de enzimas de restricción y ligasa (LIC del inglés *Ligase Independent Cloning*) mediante la técnica de fusión por PCR (Geiser *et al.*, 2001). En la **Tabla 3.5** se recogen los extremos que deben contener los oligonucleótidos para poder clonar cualquier gen en estos vectores mediante el sistema LIC. Para validar el uso de estos plásmidos como vectores eficaces para la expresión de proteínas recombinantes en *E. coli*, se preparó una batería de proteínas de fusión (**Tabla 3.6**). Todas estas proteínas se produjeron en la cepa de *E. coli* BL21(DE3), creciendo las células en medio LB-Km hasta alcanzar una DO<sub>600</sub> entre 0,4-0,6. La sobreexpresión de proteína se indujo con IPTG (0,4 mM) y posteriormente los cultivos se incubaron durante 20 horas aproximadamente a 16 °C y con agitación. Por último, las células se recogieron mediante centrifugación a 4.500 rpm (IEC CL40 Centrifuge, Thermo Scientific) durante 20 minutos a 4 °C y se conservaron a -80 °C en espera de ser procesadas.



**Figura 3.13.** Set de plásmidos diseñado en base al uso de los módulos lectina N-terminales como etiquetas de fusión. El vector **A** incorpora el módulo LSL<sub>150</sub> y el vector **B**, el módulo LBL<sub>152</sub>. En ambos casos, los módulos lectina se encuentran clonados entre las dianas *NcoI* y *EcoRI* e incorporan un *linker* y la secuencia de reconocimiento de la proteasa TEV. El vector pKLSLl es equivalente al vector pKLSLl, pero el módulo lectina incorpora las mutaciones que activan el sitio  $\alpha$ . El vector **C**, por otro lado, presenta una versión de LSL<sub>150</sub> con una cola de histidinas en el extremo N-terminal. La secuencia *linker* y el sitio de reconocimiento de la proteasa TEV se mantienen.

Número (Sentido)	Gen que se amplifica	Vector de clonación	Secuencia 5'→3'	Vector
1 (D)	<i>lsl<sub>150</sub></i>	pET-28a(+)	gatg <b>ccatgg</b> caATGACCGACATC	pKLSLt**
21 (R)	<i>lsl<sub>150</sub>- linker+TEV</i>	pET-28a(+)	<b>cgaattc</b> gccctggaaatacaggttttc agaagagcttgcATCCTCGAAGAGGAAG	pKLSLt**
3 (D)	<i>lbl<sub>152</sub></i>	pKLSLt *	gatg <b>ccatgg</b> CAAACGAGTACAATCCCCCTCTTG	pKLBLt
22 (R)	<i>lbl<sub>152</sub></i>	pKLSLt *	ctag <b>ctagc</b> GTCTGAATAAGATAGTGAAATATTG	pKLBLt
23 (D)	<i>lsl<sub>150</sub>- linker+TEV</i>	pET-28a(+)	<i>gggcagcagccatcatcatcatcac</i> ATGACCGACATCTACATCCCGCCGG	pKHLSLt
24 (R)	<i>lsl<sub>150</sub>- linker+TEV</i>	pET-28a(+)	<i>ccaccagtcatgctagccatatg</i> GCCCTGGAAATACAGGTTTTC	pKHLSLt

**Tabla 3.4.** Secuencias de los oligonucleótidos utilizados para la construcción de los vectores de expresión basados en las lectinas LSL<sub>150</sub> y LBL<sub>152</sub>. En *minúscula* y *negrita* aparecen los sitios de reconocimiento de las enzimas de restricción. En *minúscula* y *cursiva* aparecen los sitios que hibridan directamente con el plásmido para la clonación independiente de enzimas de restricción y ligasa. En *mayúscula* se indican las regiones que hibridan con el inserto. Las letras D y R indican el sentido de lectura del oligonucleótido: D significa sentido directo (5'→3') y R, sentido reverso (3'→5' de la hebra codificante).

\* Se utilizó un vector pKLSLt en el que se había generado un sitio de restricción *NheI*.

\*\* El vector pKELSLt se construyó sobre el vector pKLSLt, llevando a cabo la mutagénesis dirigida con los pares de oligonucleótidos M1/M2, M3/M4 y M5/M6.

La purificación de las proteínas de fusión se llevó a cabo mediante cromatografía de afinidad con una matriz de Sepharose® 4B (Sigma-Aldrich), siguiendo el mismo protocolo descrito en la sección 2.1 de *Materiales y Métodos*. Las fracciones que contenían la proteína de fusión fueron concentradas hasta un volumen menor de 5 mL. Finalmente, la proteína concentrada se terminó de purificar mediante cromatografía de exclusión por tamaño, utilizando las columnas HiLoad 16/600 Superdex 75 ó 200 Prep Grade (GE Healthcare Life Sciences) en función de la masa molecular de la proteína de fusión. La concentración final se determinó espectroscópicamente mediante los coeficientes de absorción  $E^{0,1\%}$  (1 cm, 280 nm) calculados teóricamente con el programa ProtParam (ExPASy) para cada proteína.

La eliminación de la etiqueta de fusión mediante la digestión con la proteasa TEV se llevó a cabo de dos formas distintas: tanto después de haber purificado la proteína de fusión como durante el primer paso cromatográfico. En ambos casos, el digerido se cargó en una columna de afinidad a Ni<sup>2+</sup> (IMAC) para eliminar la proteasa TEV del medio. En este trabajo se empleó una proteasa producida recombinantemente con una cola de histidinas en el extremo N-terminal que es activa en el tampón TSA y que actúa eficientemente en una



relación 1:80 (proteasa:proteína) durante un tiempo estimado de 20 h a 4 °C. Por otro lado, en el caso de llevar a cabo la digestión en columna, esto es, cuando la proteína de fusión está inmovilizada en la matriz de Sepharose® 4B, se acopló en serie una columna de afinidad a Ni<sup>2+</sup> para que la proteína de interés eluyese completamente limpia y sin contaminación de lactosa y de lectina (etiqueta de fusión), puesto que ésta última permanece unida a la resina de Sepharose® 4B.

	Extremo de los oligonucleótidos (5'→3')
<b>Directo</b>	TGAAAACCTGTATTTCCAGGGC <b>atg</b> ...
<b>Reverso</b>	TCAGCTTCCTTTTCGGGCTTTGTTA <b>stop</b> ...

**Tabla 3.5.** Extremos que se utilizaron en el diseño de los oligonucleótidos para la clonación de genes en los vectores de expresión pKLSlt, pKELSLt, pKLBLt y pKHLSlt mediante el método no dependiente de enzimas de restricción y ligasa (LIC). Mediante esta metodología, los genes quedan fusionados a los módulos lectina por el extremo N-terminal con la secuencia *linker* y la zona de reconocimiento de la proteasa TEV de por medio.

## 7.2. Estudios de inmovilización sobre agarosa de proteínas fusionadas a LSLt

Los estudios de inmovilización mediante el sistema LSLt sobre agarosa se llevaron a cabo con las proteínas de fusión LSLt-BTL2 y LSLt-EGFP. Ambas proteínas fueron incubadas con 1 g seco tanto de agarosa 4B como 4BCL a dos temperaturas (4 y 25 °C) durante diferentes tiempos. En el caso de LSLt-BTL2, se utilizaron 10 mL de LSLt-BTL2 a 3 ó 40 U/mL; mientras que para LSLt-EGFP se emplearon 10 mL a una concentración de 1 mg/mL. El seguimiento del proceso de inmovilización se determinó midiendo la actividad lipasa para LSLt-BTL2 (Fernández-Lorente *et al.*, 2008) o concentración de proteína para LSLt-EGFP (Bradford, 1976) del sobrenadante y de la suspensión de alícuotas centrifugadas, tomadas a distintos tiempos de incubación.

La reversión de la inmovilización por lactosa también fue estudiada. Para tal fin, 0,1 g de agarosa con LSLt-BTL2 inmovilizada fueron incubados durante 30 minutos a pH 7,0 y 25 °C con 1 mL de lactosa en solución a diferentes concentraciones (0-400 mM). Posteriormente, se midió la actividad lipasa de los sobrenadantes.

Por último, se hicieron estudios de estabilidad del dominio LSLt como agente para la inmovilización de proteínas. Muestras de LSLt-BTL2 inmovilizada en agarosa 4BCL a 25 °C durante 18 horas se incubaron en diferentes condiciones de postinmovilización. Cada

muestra (0,1 g) se incubó durante 1 hora con 1 mL de diferentes tampones (50 mM acetato de sodio pH 5,0; 50 mM fosfato de sodio pH 7,0; 50 mM carbonato de sodio pH 10,0) a distintas temperaturas (25, 40 y 65 °C) con diferentes concentraciones de acetonitrilo (20, 40 y 60%) y en presencia de 0,2% Triton X-100 a pH 7,0. La actividad lipasa fue medida tanto en los sobrenadantes como en las suspensiones.

<b>Proteína modelo</b>	<b>Organismo de origen</b>	<b>Vector de expresión</b>	<b>Método de clonación</b>	<b>Referencia bibliográfica</b>
<b>EGFP</b>	<i>Aequorea victoria</i>	pKLSLt	<i>EcoRI-NotI</i>	(Tsien, 1998)
<b>EGFP</b>	<i>Aequorea victoria</i>	pKLBLt	LIC	(Tsien, 1998)
<b>EGFP</b>	<i>Aequorea victoria</i>	pKELSLt	LIC	(Tsien, 1998)
<b>Cpl-7</b>	Bacteriófago Cp-7	pKLSLt	<i>EcoRI-HindIII</i>	(Garcia <i>et al.</i> , 1990)
<b>SMD</b>	<i>Arcanobacterium haemolyticum</i>	pKLSLt	<i>EcoRI-HindIII</i>	(McNamara <i>et al.</i> , 1995)
<b>PLD</b>	<i>Corynebacterium pseudotuberculosis</i>	pKLSLt	<i>EcoRI-HindIII</i>	(McNamara <i>et al.</i> , 1995)
<b>Bgl</b>	<i>Lactobacillus plantarum</i>	pKLSLt	LIC	(Acebrón <i>et al.</i> , 2009)
<b>PAD</b>	<i>Lactobacillus plantarum</i>	pKLSLt	LIC	(Rodríguez <i>et al.</i> , 2010)
<b>BTL2</b>	<i>Geobacillus thermocatenolatus</i>	pKLSLt	LIC	(Carrasco-López <i>et al.</i> , 2008)
<b>BTL2</b>	<i>Geobacillus thermocatenolatus</i>	pKLBLt	LIC	(Carrasco-López <i>et al.</i> , 2008)
<b>Proteasa TEV</b>	Virus del grabado del tabaco (TEV)	pKLSLt *	<i>EcoRI-XhoI</i>	(Carrington y Dougherty, 1987)
<b>SSL</b>	<i>Sclerotinia sclerotiorum</i>	pKHLSLt	LIC	(Amselem <i>et al.</i> , 2011)

**Tabla 3.6.** Resumen de las proteínas que se han clonado en el set de vectores de expresión generados a partir de los módulos lectina N-terminales. Se especifica la fuente de origen de la proteína, así como el método elegido para la clonación de los respectivos genes.

\* Hace referencia a una versión del vector pKLSLt que presenta una mutación en el sitio de reconocimiento de la proteasa TEV con el fin de evitar la autoescisión.

### 7.3. Microscopía confocal láser de barrido

La microscopía confocal láser de barrido (CLSM del inglés *Confocal Laser Scanning Microscopy*) es una técnica para obtener imágenes ópticas de alta resolución (Paddock, 1999). Permite adquirir imágenes enfocadas a diferentes planos focales (seccionamiento óptico) que, mediante un posterior análisis informático, consiguen reconstruir un objeto complejo

en tres dimensiones. Entre sus variadas aplicaciones destaca la visualización de los perfiles de superficies de objetos opacos.

Con la intención de estudiar la distribución de la proteína de fusión LSLt-EGFP una vez inmovilizada en partículas de agarosa, se utilizó CLSM para visualizar cómo se encontraban las moléculas de proteína en la superficie de partículas de agarosa 4B y 4BCL. En primer lugar, LSLt-EGFP fue rápidamente inmovilizado a 4 °C y después de 1 minuto, la suspensión se filtró para quitar toda la proteína no unida. Las partículas que contenían la proteína se diluyeron en una proporción 1:10 (partícula:tampón) en un tampón acuoso y su fluorescencia se registró a diferentes intervalos de tiempo. La monitorización se llevó a cabo con un microscopio confocal láser Leica TCS SP5. El modo de barrido secuencial se usó para evitar diafonía entre los canales. Las imágenes fueron procesadas con la aplicación LAS AF Leica y el programa Adobe Photoshop CS2 (Adobe Systems Inc.).

#### **7.4. Caracterización de la unión de lectinas a nanopartículas magnéticas**

La tecnología basada en el uso de nanopartículas<sup>28</sup> como soporte para inmovilizar proteínas fusionadas a módulos lectina en posición N-terminal fue desarrollada en colaboración con la Dra. M<sup>a</sup> del Puerto Morales Herrero del Departamento de Materiales Bioinspirados y Biomateriales del Instituto de Ciencia de Materiales de Madrid (ICMM, CSIC). La preparación y caracterización fisicoquímica de las nanopartículas magnéticas recubiertas de agarosa mediante dos procedimientos distintos y con diferentes diámetros de partícula (6, 8 y 12 nm) se llevó a cabo íntegramente en las instalaciones del ICMM.

##### **7.4.1. Estudio de la unión de lectinas a nanopartículas magnéticas recubiertas de agarosa (NPA)**

Los ensayos de unión de las nanopartículas magnéticas recubiertas de agarosa (NPA) se llevaron a cabo en un volumen final de 50-70 µL en tampón TSA, utilizando diferentes cantidades de lectina (5 y 10 µg). Las NPA se incubaron durante 2 horas con la proteína a temperatura ambiente. Posteriormente, con ayuda de un imán, se separaron los sobrenadantes de las NPA y se analizaron mediante SDS-PAGE.

Además de la unión, también se estudió la reversión de ésta en presencia de lactosa. Una vez obtenidos los complejos NPA:lectina, éstos fueron incubados en una solución 0,1 M lactosa durante 1 hora a temperatura ambiente. Después de retirar las NPA con un imán, los sobrenadantes se analizaron mediante SDS-PAGE.

Todos los ensayos se acompañaron de controles que incluían nanopartículas magnéticas sin recubrimiento, recubiertas con ácido dimercaptosuccínico (DMSA) (con carga neta negativa) y recubiertas con aminodextrano (con carga neta positiva).

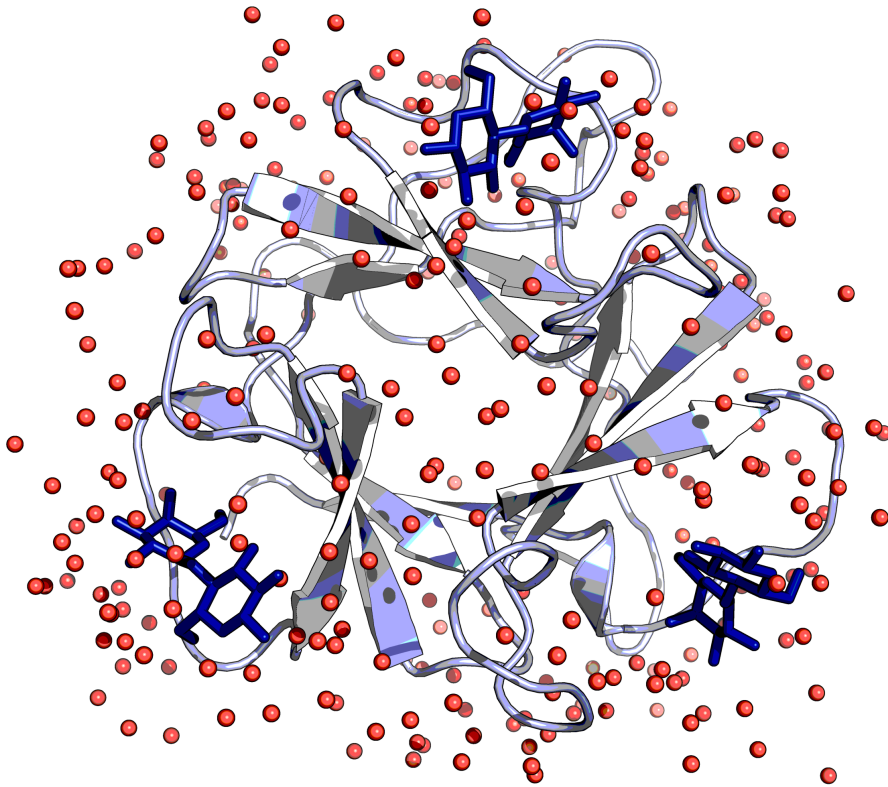
#### 7.4.2. Orientación y actividad de proteínas inmovilizadas a nanopartículas magnéticas recubiertas de agarosa (NPA)

La direccionalidad de la interacción entre las proteínas de fusión y las nanopartículas NPA se estudió mediante dos aproximaciones experimentales. Con respecto a LSLt-EGFP, se prepararon complejos incubando la proteína (100  $\mu$ g) con NPA (750  $\mu$ g) (2:5) en un volumen final de 200  $\mu$ L. Después de 2 horas de incubación, los complejos se lavaron con tampón TSA para eliminar la fracción de proteína no unida. El complejo NPA:lectina se sometió a digestión con la proteasa TEV durante 20 horas aproximadamente a 4 °C y con agitación. Mediante SDS-PAGE se estudió la cinética de la digestión, cargando los correspondientes sobrenadantes a diferentes tiempos. La fracción de proteína total unida a las nanopartículas también se analizó electroforéticamente, incubando previamente el complejo NPA:lectina con 0,1 M lactosa en tampón TSA.

Como segunda aproximación, se midió la actividad lipasa de LSLt-BTL2 inmovilizada sobre NPA. El complejo NPA:LSLt-BTL2 se preparó a partir de 300  $\mu$ g de proteína de fusión con 750  $\mu$ g de NPA (2:5) en un volumen final de 200  $\mu$ L. Después de 2 horas de incubación, las nanopartículas se lavaron con tampón TSA para retirar la fracción no unida. El ensayo de actividad lipasa (Fernández-Lorente *et al.*, 2008) sirvió para evaluar la actividad tanto de los complejos NPA:LSLt-BTL2 como de LSLt-BTL2 después de revertirse la unión con 0,1 M lactosa en tampón TSA.

#### 7.4.3. NPA como un método para limpiar un medio de reacción

Para llevar a cabo este ensayo, 100  $\mu$ g de las proteínas de fusión LSLt-EGFP y LSLt-Cpl-7 se incubaron durante ~20 horas a 4 °C con la endoproteasa TEV, producida como proteína de fusión en el vector pKLSLt (LSLt-TEV), en una relación 1:80. Los digeridos se incubaron con 300  $\mu$ g de NPA (1:3) durante 2 h a temperatura ambiente. Las NPA se separaron del sobrenadante con ayuda de un imán para ser incubadas en una solución con 0,1 M lactosa en tampón TSA durante 1 hora a temperatura ambiente. Finalmente, las NPA se retiraron con el imán de nuevo y se llevó a cabo en análisis de todos los pasos mediante SDS-PAGE.



## **RESULTADOS Y DISCUSIÓN**

**Obtención y caracterización  
de lectinas fúngicas**



## IV

## RESULTADOS Y DISCUSIÓN

## Obtención y caracterización de lectinas fúngicas

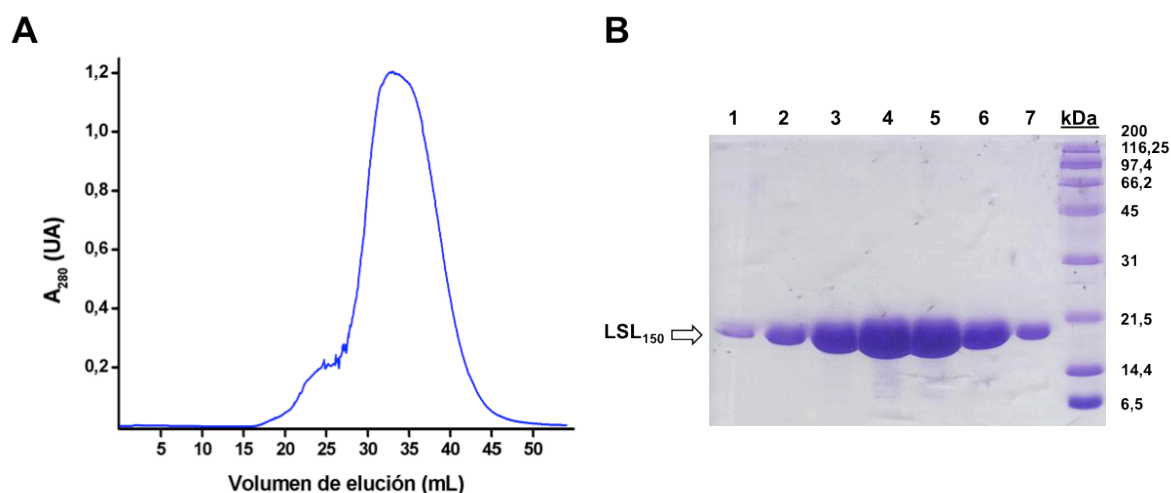
## 1. El módulo lectina N-terminal de la proteína LSLa

1.1. El módulo lectina N-terminal de la proteína LSLa de *Laetiporus sulphureus* (LSL<sub>150</sub>) es una unidad de plegamiento independiente

El módulo lectina N-terminal de la proteína LSLa de *L. sulphureus* (LSL<sub>150</sub>) se purificó a homogeneidad en un único paso cromatográfico, siguiendo el protocolo que se ha descrito en el apartado 2.2 de *Materiales y Métodos*. La proteína eluyó de la columna de Sepharose® 4B empleando el tampón TSAL, que contiene lactosa a una concentración final de 0,2 M. Tal y como revela el análisis mediante SDS-PAGE, la proteína se encuentra pura (**Figura 4.1**). La simple observación de que LSL<sub>150</sub> queda retenida en la matriz cromatográfica y que se disocia de ella por adición de lactosa apunta a que la lectina mantiene intacta su capacidad de unión a carbohidratos y, por lo tanto, está bien plegada. El rendimiento en la producción del módulo lectina en las células de *E. coli* BL21(DE3) fue aproximadamente de ~80 mg/L de cultivo. Este resultado es consistente con el obtenido previamente en la producción del mutante de delección LSLa-D1 en células de *E. coli* NovaBlue(DE3) (~60 mg/L), que contiene los primeros 187 aminoácidos de la secuencia de LSLa (Tateno y Goldstein, 2003). Estos resultados indican, por un lado, que la secuencia de nucleótidos que codifica el módulo lectina N-terminal de LSLa exhibe una alta capacidad de traducción y, por otro, que *E. coli* proporciona un entorno celular adecuado para un eficiente y correcto plegamiento de un módulo lectina de origen fúngico.

Si se compara el rendimiento de producción de LSL<sub>150</sub> con el obtenido para otras lectinas fúngicas con estructura trébol  $\beta$  expresadas recombinantemente, se observa que en general todas ellas se producen con elevados rendimientos. Por ejemplo, las lectinas rCNL de *C. nebularis* y CCL2 de *C. cinerea* también presentan rendimientos de producción elevados, siendo incluso superiores a los de LSL<sub>150</sub>: ~190 mg/L para el caso de rCNL en cé-

lulas de *E. coli* BL21(DE3) (Pohleven *et al.*, 2012). Sin embargo, cuando la lectina es un módulo que forma parte de una proteína más grande, la proteína completa expresada recombinantemente tiene asociados rendimientos más bajos. En el caso de la lectina rPSL de *P. squamosus*, por ejemplo, el rendimiento de producción en células de *E. coli* Nova-Blue(DE3) se encuentra en torno a los 4-7 mg/L (Tateno *et al.*, 2004). Este comportamiento también es observado en la producción recombinante de la proteína completa rLSLa, cuyo rendimiento final no supera los 4 mg/L de cultivo (Tateno y Goldstein, 2003).



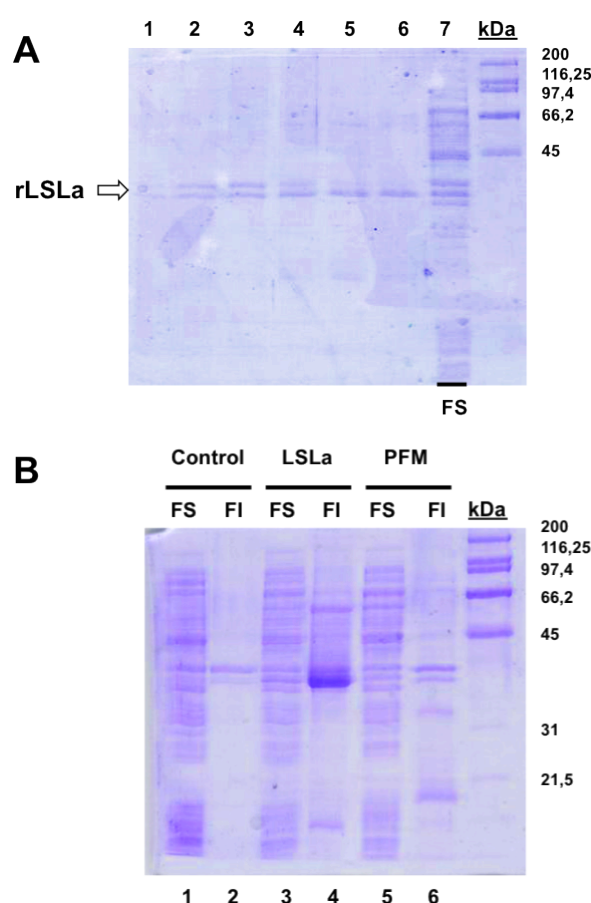
**Figura 4.1.** Purificación del módulo lectina N-terminal LSL<sub>150</sub>. **A.** Perfil de elución de la purificación de LSL<sub>150</sub> mediante cromatografía de afinidad a Sepharose® 4B. **B.** SDS-PAGE (15% de poliacrilamida) de la purificación de LSL<sub>150</sub> eluido con 0,2 M lactosa de la matriz cromatográfica. Los carriles 1-7 se corresponden con el máximo observado en el perfil de elución.

## 1.2. El módulo LSL<sub>150</sub> actúa como un agente solubilizante natural

La comparación de los rendimientos de producción del módulo LSL<sub>150</sub> y el de la proteína LSLa completa revela un comportamiento operativamente similar al modo de actuación de muchas etiquetas de solubilidad empleadas en sistemas de expresión de proteínas recombinantes (Malhotra, 2009). En este sentido, también es destacable el hecho de que no existan apenas contactos entre el módulo lectina N-terminal y el módulo C-terminal formador de poros (PFM) dentro de la estructura de la proteína completa LSLa. Esta característica, de hecho, acentúa la independencia de ambos módulos y, por tanto, explicaría que LSL<sub>150</sub> se comporte como una unidad independiente de plegamiento. A ello hay que sumar, como se verá más adelante, que LSL<sub>150</sub> adquiere la misma conformación observada en la estructura nativa del módulo lectina de LSLa y que mantiene intacta la capacidad de unir carbohidratos, es decir, de actuar como una lectina. Estas observaciones nos llevaron a la hipótesis de trabajo que considera al módulo lectina como un potenciador de la solubilidad del módulo



C-terminal que lo acompaña. Aunque la comprobación de esta hipótesis requeriría de un análisis *in vivo* en el hongo productor, pueden obtenerse evidencias indirectas empleando *E. coli* como sistema de expresión. De este modo, se expresaron de forma recombinante tanto la proteína LSLa como el módulo PFM en células de *E. coli* BL21(DE3). Los resultados obtenidos (**Figuras 4.2A y 4.2B**) permiten hacer un par de consideraciones: en primer lugar, se produce recombinantemente LSLa (rLSLa con una masa molecular de 35 kDa), la cual se pudo purificar utilizando el mismo protocolo seguido en la obtención de LSL<sub>150</sub>; en segundo lugar, PFM se produjo exclusivamente en la fracción insoluble (18 kDa). Aunque bien es cierto que la mayor parte de rLSLa se encuentra en la fracción insoluble, los resultados obtenidos apuntan a que en *E. coli* el módulo lectina actuaría como un potenciador de la solubilidad del módulo formador de poros, apoyando por ello la hipótesis de trabajo antes considerada.



**Figura 4.2.** Comprobación del comportamiento de LSL<sub>150</sub> como un potenciador natural de la solubilidad. **A.** SDS-PAGE (12% de poliacrilamida) de la purificación de rLSLa. Los carriles 1-6 describen el máximo del perfil de elución; mientras que el carril 7, se corresponde con la fracción soluble (FS) del extracto celular de las células de *E. coli* BL21(DE3) transformadas con el plásmido pKLSLa. **B.** Comparación de la solubilidad de rLSLa y rPFM. FS se corresponde con la fracción soluble del extracto celular y FI, con la fracción insoluble del mismo. Como control se han utilizado células de *E. coli* BL21(DE3) transformadas con el vector pET-28a(+), que no sobreexpresa ninguna proteína bajo inducción con IPTG.

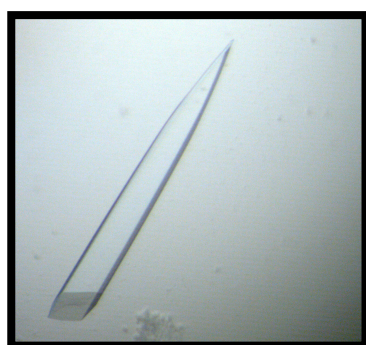
Es interesante destacar por otro lado que en la fracción insoluble del extracto celular de las células de *E. coli* BL21(DE3) transformadas con el plásmido pKLSLa, no sólo se observa la banda correspondiente a rLSLa, sino que se aprecia una banda más pequeña que

se correspondería con una forma truncada de la lectina. Esta banda ya había sido identificada anteriormente como la lectina rLSLa después de haber sufrido una posible digestión del extremo C-terminal con proteasas bacterianas (Tateno y Goldstein, 2003).

### 1.3. El módulo LSL<sub>150</sub> es un trébol $\beta$ que une lactosa

#### 1.3.1. Cristalización de LSL<sub>150</sub> y del complejo con lactosa

Las condiciones de cristalización de LSL<sub>150</sub> se establecieron llevando a cabo un muestreo inicial sobre una amplia variedad de condiciones de cristalización (Jancarik y Kim, 1991) por el método de difusión de vapor en gota colgante a una temperatura de 291 K. La condición optimizada para cristalizar LSL<sub>150</sub> se obtuvo mezclando 2  $\mu$ L de solución de proteína (25-30 mg/mL) en tampón TSA con 2  $\mu$ L de la siguiente solución precipitante: 12% (p/v) PEG 4000, 100 mM Tris-HCl pH 8,5 y 150 mM acetato de sodio. Los cristales (**Figura 4.3**) aparecieron en un día y crecieron hasta unas dimensiones aproximadas de 0,1×0,3×0,3 mm<sup>3</sup> en el transcurso de 3-4 días.



**Figura 4.3.** Cristal de LSL<sub>150</sub> obtenido en 12% (p/v) PEG 4000, 100 mM Tris-HCl pH 8,5 y 150 mM acetato de sodio.

LSL<sub>150</sub> fue asimismo cocrystalizada en presencia de lactosa, incubando previamente la proteína con tampón TSAL en el que la concentración final del disacárido es 0,2 M. Los cristales del complejo se obtuvieron en la misma condición de cristalización anterior, pero después de 2-3 semanas y alcanzando unas dimensiones medias de 0,2×0,3×0,5 mm<sup>3</sup>.

#### 1.3.2. Determinación estructural de LSL<sub>150</sub> y del complejo con lactosa

Los cristales del módulo lectina con y sin lactosa que se utilizaron para la resolución estructural fueron crioprotectados con una solución que contenía la condición de cristalización suplementada con un 20% (v/v) de glicerol. La toma de datos de los cristales sin ligando se llevó a cabo en la línea ID23-1 del ESRF (Grenoble, Francia) a una temperatura de

100 K y usando un detector ADSC Q315. En cuanto a los cristales con lactosa, estos se midieron en las mismas condiciones y con el mismo detector en la línea ID29 del mismo sincrotrón.

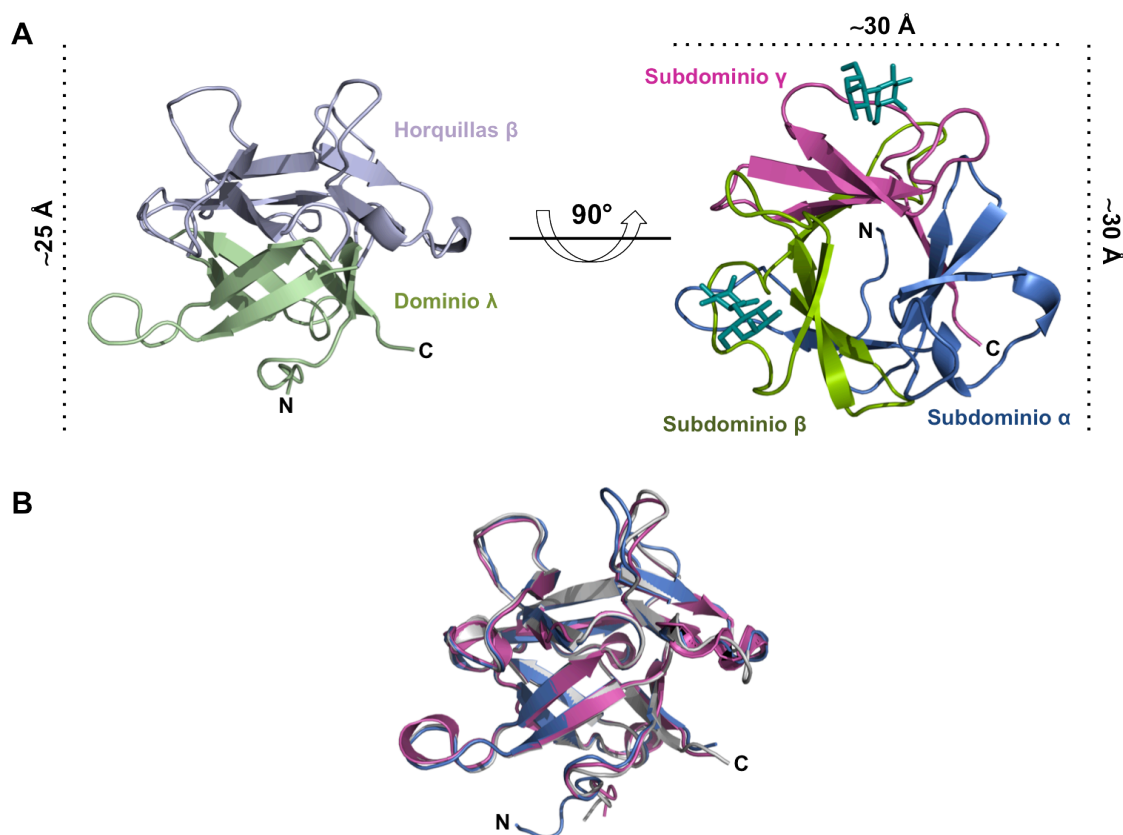
El análisis cristalográfico reveló que los cristales de LSL<sub>150</sub> contienen una sola molécula en la unidad asimétrica y pertenecen al grupo espacial  $P2_12_12_1$  con los siguientes parámetros de celda unidad:  $a = 34,58 \text{ \AA}$ ,  $b = 59,54 \text{ \AA}$  y  $c = 61,62 \text{ \AA}$ . Por otro lado, los cristales del complejo con lactosa son trigonales y pertenecen al grupo espacial  $P3_2$ , con parámetros de celda unidad:  $a = 62,11 \text{ \AA}$ ,  $b = 62,11 \text{ \AA}$  y  $c = 37,65 \text{ \AA}$ . En ambos casos, se obtuvieron resoluciones elevadas, siendo éstas de 1,47 y 1,67  $\text{\AA}$  para LSL<sub>150</sub> y LSL<sub>150</sub> en complejo con lactosa, respectivamente.

Las estructuras de la lectina con y sin ligando se obtuvieron mediante reemplazo molecular, usando como modelo las coordenadas de los 150 primeros residuos del dominio lectina N-terminal de LSLa (código PDB 1W3A), que había sido resuelta con antelación (Mancheño *et al.*, 2005). Después de varios ciclos de refinamiento, aplicando grupos TLS (del inglés *translation-libration-screw*), se obtuvieron unos modelos finales con unos valores de  $R_{work}$  y  $R_{free}$  de 15,7 y 19,4% para LSL<sub>150</sub> y 17,6 y 22,9% para el correspondiente complejo con lactosa. El resto de datos de la colección de las imágenes de difracción y de las estadísticas del refinamiento están resumidos en la **Tabla 4.1**.

Los resultados obtenidos revelan que el plegamiento de LSL<sub>150</sub> con y sin ligando es prácticamente idéntico al descrito previamente para el dominio lectina N-terminal de LSLa (**Figura 4.4**) (Mancheño *et al.*, 2005). De hecho, si se superponen las estructuras de LSL<sub>150</sub> con y sin ligando y el módulo N-terminal de LSLa, el RMSD medio de la cadena polipeptídica ( $C\alpha$ ) es de 0,7  $\text{\AA}$  (**Figura 4.4B**). De esta manera, la estructura de LSL<sub>150</sub> es la típica de un trébol  $\beta$ , con pseudosimetría de orden 3, resultado de la repetición de un motivo estructural básico que origina los tres subdominios:  $\alpha$ ,  $\beta$  y  $\gamma$ . El trébol está constituido por un pequeño barril  $\beta$  formado por seis hebras  $\beta$  antiparalelas que está coronado por un triplete de horquillas  $\beta$ , donde se localizan los residuos que participan en la unión con el ligando (apartado 2.1 de la *Introducción*) (**Figura 4.4A**). Al igual que ocurría con LSLa, LSL<sub>150</sub> presenta únicamente los sitios  $\beta$  y  $\gamma$  operativos para unir lactosa.

El análisis del empaquetamiento de los cristales de LSL<sub>150</sub> sin lactosa con el servidor PISA (Krissinel y Henrick, 2007) revela que la proteína está presente como una especie monomérica (las áreas de contacto determinadas están en el intervalo 129-396  $\text{\AA}^2$ , por debajo del 5% del área accesible de LSL<sub>150</sub> al solvente). Este resultado coincide con lo obser-

vado en la cromatografía analítica de exclusión molecular, donde aparece un único pico correspondiente a una especie de ~17 kDa y con el hecho de que LSL<sub>150</sub> no participa en los contactos entre las subunidades del homohexámero de LSLa (Mancheño *et al.*, 2005).



**Figura 4.4.** Estructura cristalográfica del módulo lectina N-terminal LSL<sub>150</sub>. **A.** Visión frontal (izquierda) de LSL<sub>150</sub> donde se representa el dominio λ (verde claro) y el triplete de horquillas β (violeta). Visión cenital (derecha) de LSL<sub>150</sub> con las moléculas de lactosa unidas (representadas en el modelo de varillas y color cian) en los sitios β y γ. Además, se muestran en diferentes colores los tres subdominios que conforman el trébol β: α (azul marino), β (verde) y γ (magenta). Las dimensiones aproximadas del trébol β se indican en los laterales y en angstrom (Å). **B.** Superposición de las estructuras tridimensionales de LSL<sub>150</sub> (azul), LSL<sub>150</sub> en complejo con lactosa (magenta) y el correspondiente dominio lectina N-terminal de la proteína LSLa (gris).

La conservación de la estructura tridimensional de LSL<sub>150</sub> y por ello de su capacidad de unión a azúcares, así como su homogeneidad y estabilidad en solución (**Figura 4.1**) demuestran que LSL<sub>150</sub> es una unidad de plegamiento independiente, que operativamente pudiera estar actuando *in vivo* como un agente solubilizante para el módulo formador de poros en membranas, además de estar implicado en el reconocimiento del receptor molecular de la toxina.

Parámetros	LSL <sub>150</sub>	LSL <sub>150</sub> :(lactosa) $\beta\gamma$
Recogida de datos y procesamiento		
Grupo espacial	$P2_12_12_1$	$P3_2$
Moléculas en la UA	1	1
Celda unidad (Å)	34,58, 59,54, 61,62	62,15, 62,15, 37,65
Longitud de onda (Å)	0,979	0,979
Resolución (Å)	34–1,47	53–1,67
Reflexiones totales	140.792	112.070
Reflexiones únicas	22.387	18.678
Compleitud	99,5 (98,4)	99,0 (98,1)
Multiplicidad	6,3 (6,0)	6,0 (6,0)
$R_{merge}$ (%)	6,4 (29,7)	9,1 (32,7)
$I/\sigma(I)$	16,9 (5,4)	12,6 (5,2)
Refinamiento		
$R_{work}/R_{free}$	0,157/0,194	0,176/0,229
RMSD de los enlaces (Å)	0,030	0,026
RMSD de los ángulos (°)	2,315	2,144
Factor $B$ medio (Å <sup>2</sup> )	8,65	5,31
Número de átomos		
Proteína	1.212	1.197
Glicerol	0	6
Lactosa	0	46
Agua	252	262
Mapa de Ramachandran		
Favorecidos (%)	96,6	98,6
Permitidos (%)	3,4	1,4
Prohibidos (%)	0	0

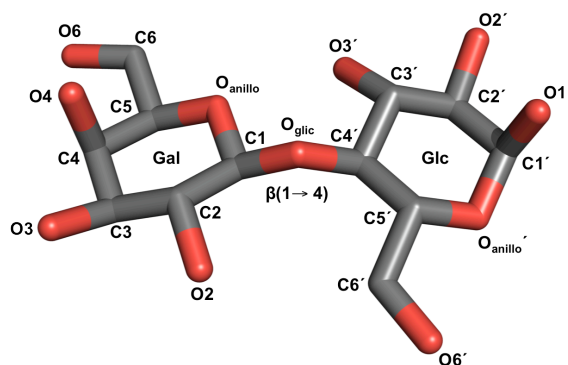
\*Los valores de la concha de mayor resolución se muestran en paréntesis.

**Tabla 4.1.** Datos de difracción y estadísticas del refinamiento de las estructuras de LSL<sub>150</sub> y LSL<sub>150</sub> en complejo con lactosa (LSL<sub>150</sub>:(lactosa) $\beta\gamma$ ).

### 1.3.3. Análisis cristalográfico de la unión a lactosa

La alta resolución de los datos cristalográficos obtenidos para el complejo de LSL<sub>150</sub> con lactosa ([LSL<sub>150</sub>:(lactosa) $\beta\gamma$ ]) (**Tabla 4.1**) ha permitido estudiar en gran detalle el modo de unión del disacárido lactosa (**Figura 4.5**) a la lectina. Asimismo, los valores de los factores  $B$  de los átomos de los azúcares y de los aminoácidos implicados en la interacción con el ligando son similares, lo que sugiere que la ocupación es completa y que, por lo tanto, todas las moléculas de proteína unen lactosa en los dos sitios operativos disponibles. Por otro lado, la excelente calidad del mapa ha facilitado la detección de cadenas laterales con dobles conformaciones, así como la identificación de las dos formas anoméricas que presenta la lactosa:  $\alpha$  (con el grupo hidroxilo del C1' del anillo de glucosa en orientación ecuatorial) y  $\beta$  (con ese mismo grupo hidroxilo en posición axial). Ambos anómeros pudieron ser modelados y refinados con ocupaciones de 0,40 y 0,60, respectivamente, en ambos sitios de unión ( $\beta$  y  $\gamma$ ). Las ocupaciones se ajustaron manualmente, buscando el mejor resul-

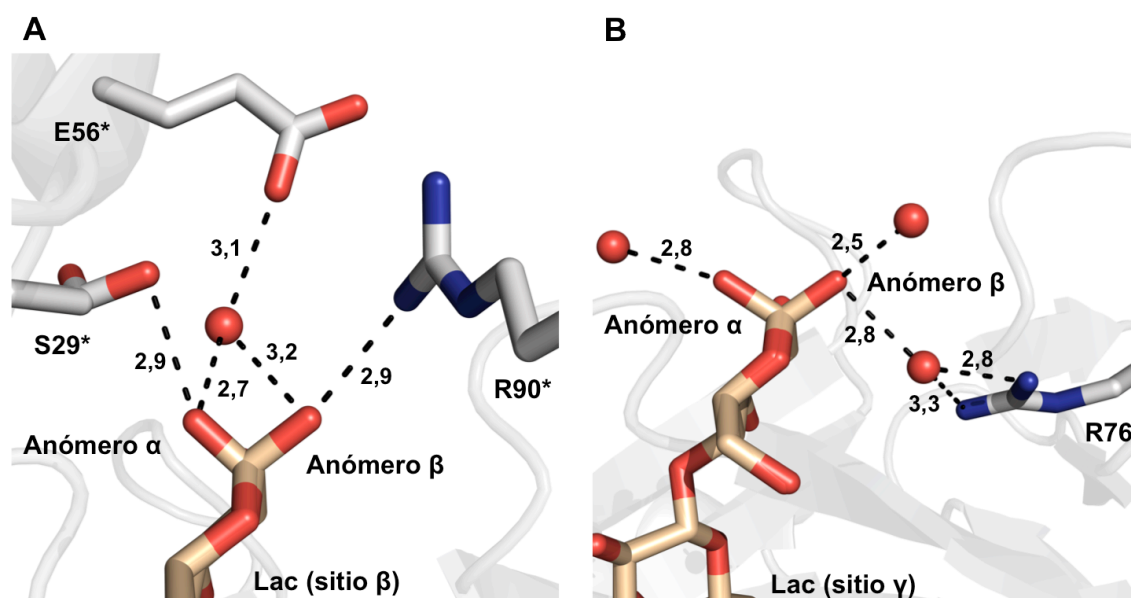
tado en función de la calidad del mapa ( $F_o - F_c$ ) en torno a la densidad del oxígeno O1' del anillo de glucosa. Teniendo en cuenta que en solución la relación  $\alpha/\beta$  para la glucosa es 0,38/0,62 (Vyas *et al.*, 1994), se puede concluir que no existe una preferencia por ninguno de los anómeros, es decir, la lectina no dispone de especificidad anomérica.



**Figura 4.5.** Estructura de la molécula de lactosa (D-galactopiranosil- $\beta(1\rightarrow4)$ -D-glucopiranososa o Gal $\beta(1\rightarrow4)$ Glc) en modelo de varillas, donde se indica la numeración de todos sus átomos (carbono en gris y oxígeno en rojo).

El estudio detallado de las interacciones permitió identificar, no obstante, interacciones específicas para cada uno de los anómeros en la estructura cristalográfica (**Figura 4.6**). Probablemente, la especificidad que se observa por un anómero u otro no tenga ninguna implicación biológica, ya que el grupo hidroxilo del carbono anomérico no está directamente implicado en interacciones con los sitios de unión de la lectina. De hecho, estas interacciones son cristalográficas, por lo que no son relevantes para el reconocimiento de potenciales ligandos. En este contexto, la lactosa presente en el sitio  $\beta$  establece diferentes enlaces de hidrógeno con residuos de moléculas de LSL<sub>150</sub> relacionadas por simetría. De esta manera, el O1' del anómero  $\alpha$  (OH del carbono C1' en posición axial) interacciona con el oxígeno de la cadena lateral de la Ser-29; mientras que en el anómero  $\beta$  (OH del carbono C1' en posición ecuatorial), lo hace con uno de los nitrógenos N $\eta$  del residuo de Arg-90. Es interesante el comportamiento de la anterior Ser-29, ya que el análisis de la densidad electrónica indica la presencia de dos confórmeros con ocupaciones parecidas a las de los anómeros de lactosa. Curiosamente, la participación de los grupos OH de residuos de Ser y Thr es infrecuente en los procesos de reconocimiento de azúcares por parte de las lectinas. La libertad rotacional que presenta el grupo OH en estos aminoácidos frente a otros grupos aceptores y dadores de enlace de hidrógeno con geometría plana hace que no contribuyan a determinar la especificidad en la unión (Weis y Drickamer, 1996). Este argumento apoya el hecho de que las interacciones caracterizadas para el oxígeno O1' del carbono anomérico son importantes en el cristal, pero no en la naturaleza. Por otro lado, la lactosa del sitio  $\gamma$

establece enlaces de hidrógeno con moléculas de agua. Además, el O1' del anómero  $\beta$  presenta más contactos ya que participa en dos enlaces de hidrógeno con moléculas de agua, estando una de ellas estabilizada con el residuo Arg-76.



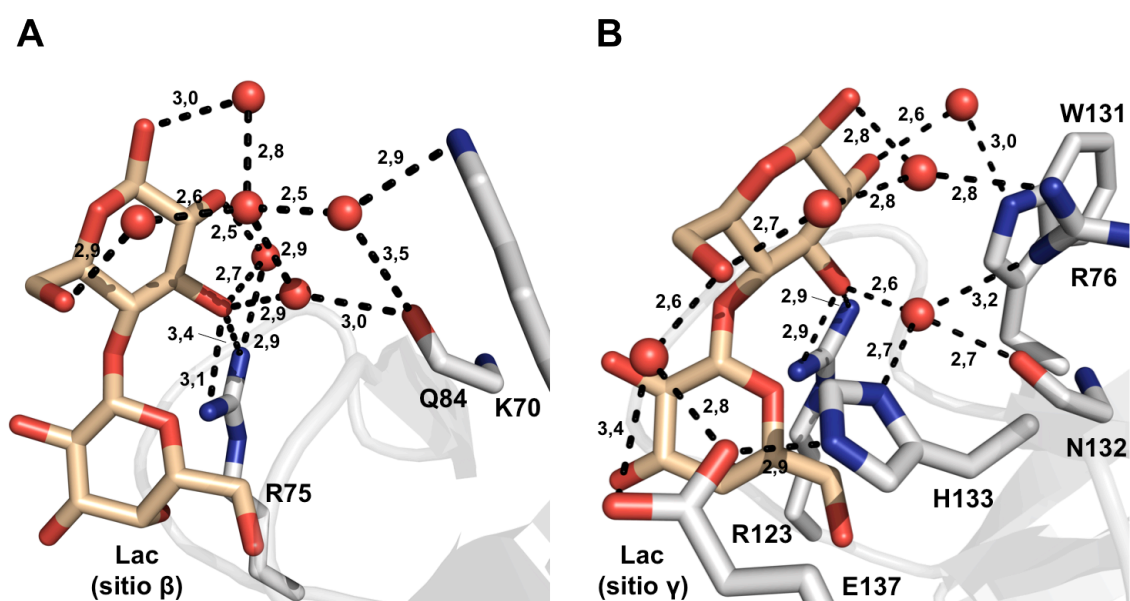
**Figura 4.6.** Vista detallada de las interacciones específicas entre los oxígenos del carbono anomérico de ambos anómeros  $\alpha$  y  $\beta$  de la lactosa y residuos de la molécula LSL<sub>150</sub> en el sitio  $\beta$  (A) y en el sitio  $\gamma$  (B). Las moléculas de agua se representan mediante *esferas* de color rojo. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno que se encuentran dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea de puntos* y las medidas de las distancias vienen dadas en *ångströms* (Å). El asterisco (\*) indica que los aminoácidos pertenecen a moléculas de proteína relacionadas por simetría.

Además de las interacciones que involucran al oxígeno O1' del carbono anomérico, se observan otras interacciones entre los oxígenos O2', O3' y O6' del residuo de glucosa y la proteína (**Figura 4.7**). Con respecto al sitio  $\beta$ , el oxígeno O2' establece un enlace de hidrógeno con una molécula de agua que, a su vez, interacciona con el oxígeno O3' y con uno de los nitrógenos N $\eta$  del grupo guanidinio de la Arg-75. Además, este átomo O3' también establece enlaces de hidrógeno directos con la cadena lateral de este aminoácido y con el oxígeno carbonílico de la cadena peptídica de la Gln-84 a través de un puente con otra molécula de agua. Esta última molécula de agua participa en una red de enlaces de hidrógeno que implica a tres moléculas de agua más, a los oxígenos O1' y O6' de la glucosa y a los residuos Lys-70 (grupo  $\epsilon$ -amino) y Gln-84 (**Figura 4.7A**).

Por otro lado, en el sitio  $\gamma$  se observa un mayor número de interacciones azúcar-proteína (**Figura 4.7B**). En primer lugar, el oxígeno O2' interacciona con una molécula de agua, que a su vez lo hace con el átomo de nitrógeno N $\epsilon$ -1 del residuo de Trp-131. El oxígeno O3' interacciona de manera similar a como lo hacía en el sitio  $\beta$ , formando enlaces de



hidrógeno con la cadena lateral de la Arg-123 y con una molécula de agua, que está estabilizada mediante interacciones con el oxígeno carbonílico de la Gln-132, uno de los nitrógenos N $\eta$  de la Arg-76 y el nitrógeno N $\delta$ -1 de la His-133. Por último, el oxígeno O6' interacciona con dos moléculas de agua. La primera de ellas (a 2,7 Å) está estabilizada por la participación en una red de moléculas de agua, mientras que la segunda (a 2,6 Å) interacciona con el grupo ácido del residuo de Glu-137. Este aminoácido ancla, a su vez, a la His-133 mediante una interacción con el nitrógeno N $\epsilon$ -2, que posiblemente ayuda a la correcta orientación del anillo imidazólico de la His-133 para enlazar con la molécula de agua citada anteriormente.



**Figura 4.7.** Vista detallada de las interacciones específicas entre el anillo de glucosa de la lactosa (se representa solamente el anómero β) y los residuos de los sitios de unión de LSL<sub>150</sub>. En **A** se muestran las interacciones del sitio β y en **B**, las del sitio γ. Las moléculas de agua se representan mediante *esferas* de color rojo. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno que se encuentran dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea de puntos* y las medidas de las distancias vienen dadas en ångströms (Å).

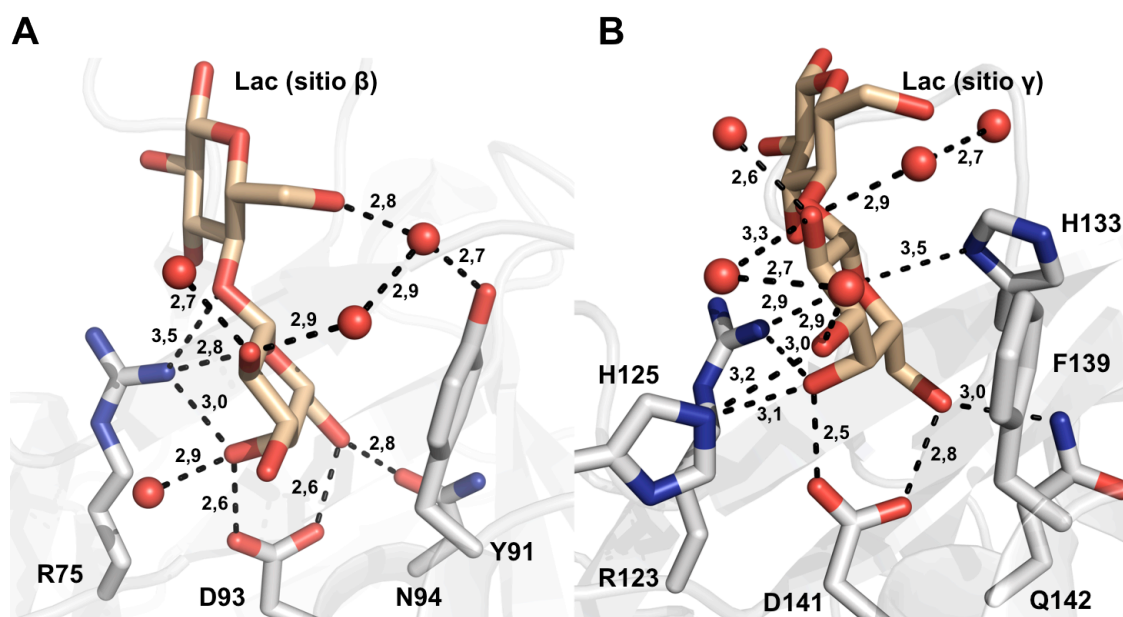
Es, sin embargo, el anillo de galactosa el que establece un mayor número de contactos con la lectina al encontrarse éste más enterrado, formando fuertes interacciones de apilamiento con residuos aromáticos: la Tyr-91 en el sitio β y la Phe-139 en el sitio γ. En ambos casos, la interacción tiene lugar a través de la cara B del anillo de galactosa, que presenta un parche hidrofóbico formado por los átomos de carbono C3, C4, C5 y C6 (Rini, 1995) (**Figura 4.8**).

Además del apilamiento con residuos aromáticos, el anillo de galactosa interacciona mediante enlaces de hidrógeno, que contribuyen a fijar aún más las moléculas de ligando en los sitios de unión (**Figura 4.8**). De este modo, en el sitio β (**Figura 4.8A**), el átomo de



oxígeno O2 de la galactosa interacciona con uno de los nitrógenos N $\eta$  de la Arg-75 y con dos moléculas de agua. Una de éstas (la que se encuentra a 2,9 Å) es estabilizada por otra molécula de agua, que está interaccionando con el oxígeno O6' de la glucosa y el oxígeno O $\eta$  de la Tyr-91. Por otro lado, el oxígeno O4 interacciona también con una molécula de agua y con el carboxilato del Asp-93. Este residuo forma también un enlace de hidrógeno con el oxígeno O6, que a su vez interacciona con el oxígeno O $\delta$ -1 de la Asn-94. Por último, la Arg-75 interacciona mediante enlace de hidrógeno con el oxígeno del enlace glicosídico y el oxígeno del anillo de galactosa.

El sitio  $\gamma$  muestra mayor número de interacciones (**Figura 4.8B**). En este caso, los oxígenos O2 y O3 de la galactosa forman parte de una red de puentes con moléculas de agua. El oxígeno O4, por su parte, interacciona con uno de los átomos N $\eta$  de la Arg-123, el Asp-141 (equivalente al Asp-93 del sitio  $\beta$ ) y con el nitrógeno N $\epsilon$ -2 de la His-125, que participa en otro enlace de hidrógeno con el oxígeno O3. El Asp-141 también está implicado en un enlace de hidrógeno con el oxígeno O6, que interacciona al mismo tiempo con el nitrógeno N $\epsilon$ -2 de la Gln-142. Por último, la Arg-123 y la His-133 estabilizan a la molécula de lactosa reconociendo el oxígeno del anillo de galactosa.



**Figura 4.8.** Vista detallada de las interacciones específicas entre el anillo de galactosa de la lactosa (se representa solamente el anómero  $\beta$ ) y los residuos de los sitios de unión de LSL<sub>150</sub>. En **A** se muestran las interacciones del sitio  $\beta$  y en **B**, las del sitio  $\gamma$ . Las moléculas de agua se representan mediante *esferas* de color rojo. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno que se encuentran dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea de puntos* y las medidas de las distancias vienen dadas en *ångströms* (Å).

Finalmente, es interesante mencionar que aunque el subdominio  $\alpha$  conserva esencialmente la misma arquitectura que la descrita en los sitios  $\beta$  y  $\gamma$ , el sitio  $\alpha$  no es capaz de unir moléculas de azúcar bajo nuestras condiciones experimentales. En este sentido, la presencia del residuo Ala-43 en lugar de un residuo aromático equivalente a los residuos Tyr-91 (sitio  $\beta$ ) y Phe-137 (sitio  $\gamma$ ) puede ser la razón principal que explicaría la ausencia de unión en este subdominio (Mancheño *et al.*, 2005).

#### 1.3.4. Comparación del modo de unión del ligando en la familia de los tréboles $\beta$

Como se comentó anteriormente en la sección 2.2 del apartado de *Introducción*, el plegamiento de tipo trébol  $\beta$  está fuertemente conservado a pesar de la pronunciada divergencia evolutiva observada en las secuencias de aminoácidos de estas proteínas. Esta divergencia ha contribuido a que muchas de las lectinas con estructura trébol  $\beta$  dispongan de mecanismos de unión específicos. No obstante, el análisis detallado de las estructuras encontradas en el *Protein Data Bank* (PDB) tras una búsqueda con las palabras clave  *$\beta$ -trefoil* y *galactose* (o *lactose*) ha revelado determinantes fundamentales del reconocimiento de azúcares, algunos de los cuales aparecen en LSL<sub>150</sub>.

A la vista de la **Tabla 4.2**, se observa que siempre existe un residuo aromático que participa en interacciones de apilamiento con el anillo de galactosa. En la mayoría de las lectinas, el aminoácido responsable de la interacción apolar es un triptófano. Incluso cuando la galactosa presenta alguna modificación covalente en el anillo, esta interacción se mantiene, tal y como ocurre en el caso de la N-acetilgalactosamina-4-sulfato (GalNAc-4-SO<sub>3</sub>) (Liu *et al.*, 2000). Esta interacción se complementa con la formación de enlaces de hidrógeno. Generalmente, el oxígeno O4 de la galactosa interacciona direccionalmente con un residuo de ácido aspártico. En algunos casos minoritarios, esa interacción es llevada a cabo por una asparagina, como en el caso del dominio del receptor de manosa rico en cisteínas (PDB 1DQO) (Liu *et al.*, 2000). Finalmente, aunque no aparece en la estructura del complejo de LSL<sub>150</sub> con lactosa, muchas lectinas de esta familia muestran una interacción específica que implica al oxígeno O3 del anillo de galactosa. En esta interacción hay más variabilidad en el aminoácido que interviene, aunque predominan ácidos aspárticos (el mismo del enlace de hidrógeno con el O4) y asparaginas.

Como se puede observar, la mayoría de los enlaces de hidrógeno se establecen con aminoácidos del sitio de unión, cuyos grupos dadores o aceptores del enlace presentan una

geometría plana y fija, es decir, cuya libertad conformacional está limitada. De hecho, como se indicó anteriormente, es muy poco frecuente la presencia de grupos OH procedentes de Ser o Thr que estén involucrados en el reconocimiento. Presumiblemente, el coste entrópico para fijar una conformación de estos grupos hidroxilo durante la unión del ligando es muy elevado, lo que resulta en su baja participación (Weis y Drickamer, 1996). Por otro lado, las cadenas laterales de geometría plana tienen la ventaja de poder formar más de un enlace con los grupos OH de las moléculas de azúcar simultáneamente porque tienen una geometría óptima para el enlace de hidrógeno (Weis y Drickamer, 1996). Por ejemplo, los residuos de Asp-93 y Asp-141 en los sitios  $\beta$  y  $\gamma$  de LSL<sub>150</sub>, respectivamente, fijan al mismo tiempo los OH de las posiciones C4 y C6 del anillo de galactosa (separados entre sí 3,6 Å) a través de los dos átomos de oxígeno de la cadena lateral (separados entre sí 2,6 Å). Por último, cabe indicar que los aminoácidos con geometría plana que interaccionan directamente con los azúcares son claves para llevar a cabo un reconocimiento específico y discriminar entre diferentes ligandos.

Código PDB	Especificidad	Sitio de unión	Aromático	Ligando-O4	Ligando-O3
1HWO	Gal	1 $\alpha$	Trp39	—	—
1IT0	Gal	$\alpha$	Trp340	Asp325	Asp325, His343, Asn347
	Gal	$\gamma$	Trp423	Asp408	Asp408, Asn430
1KNM	Gal	$\alpha$	Ttrp34	Asp19	Asp19, His37, Asn41
	Gal	$\gamma$	Tyr117	Asp102	Asp102, Asn124
1PUU	Gal	1 $\alpha$	Trp38	Asp23	Lys41
	Gal	2 $\gamma$	Tyr249	Asp235	Asn256
1RZO	Gal	1 $\alpha$	Trp2037	Asp2022	Lys2040
	Gal	1 $\gamma$	Trp2093	—	—
1YF8	Gal	1 $\alpha$	Trp34	—	Asp19
	Gal	1 $\beta$	Phe75	Val81	—
	Gal	2 $\gamma$	Tyr241	—	Asp229
2AAI	Gal	1 $\alpha$	Trp37	Asp22	Asp22, Asn46
	Gal	2 $\gamma$	Tyr248	Asp234	Asp234, His251
2ZQN	Gal	$\alpha$	Trp161	Asp146	Asp146, Lys164, Asn171
	Gal	$\gamma$	Trp245	Asp230	Asp230, Asn252, His248
2Z48	GalNAc	1 $\alpha$	Tyr36	Asp23	Asp39
	GalNAc	1 $\gamma$	Tyr134	Asp121	Gln137
	GalNAc	2 $\alpha$	Tyr181	Asp168	Gln184
	GalNAc	2 $\beta$	Tyr222	Asp209	—
	GalNAc	2 $\gamma$	Trp269	Asp256	Asp272
1DQO	GalNAc-4-SO <sub>4</sub>	$\gamma$	Trp117	Asn102	—

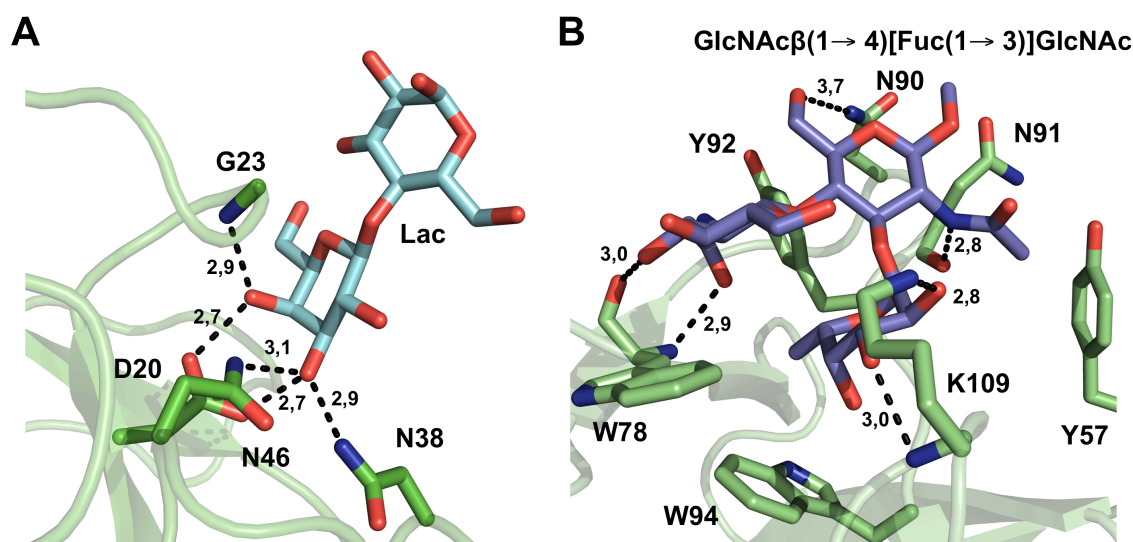
\*Solo se han analizado los residuos que interaccionan directamente con el anillo de galactosa.

**Tabla 4.2.** Comparación de los sitios operativos de unión a azúcar entre algunos miembros de la familia estructural de los tréboles  $\beta$ .

Lectina	Especificidad	Sitio de unión	Aromático	Ligando-O4	Ligando-O3
MOA	Gal $\alpha$ (1 $\rightarrow$ 3)Gal	$\alpha$	Trp35	Asp20	—
	Gal $\alpha$ (1 $\rightarrow$ 3)Gal	$\beta$	Trp87	Asp72	—
	Gal $\alpha$ (1 $\rightarrow$ 3)Gal	$\gamma$	Trp138	Asp123	—
SSA	GalNAc	$\alpha$	Tyr37	Asn22	Asn22, Asn46
PSL	Neu5Ac $\alpha$ (2 $\rightarrow$ 6)Gal	$\beta + \alpha$	Tyr87	Asp72	Asp72, Asn94
CNL	LacdiNAc	$\alpha$	—	Asp20	Asp20, Asn38, Asn46
CCL2	Fuc $\alpha$ (1 $\rightarrow$ 3)GlcNAc	Interfase $\beta$ - $\gamma$	—	—	—
RSA	GalNAc	$\alpha$	Tyr37	Asp22	Asp22, His40, Asn44
	GalNAc	$\gamma$	Tyr127	Asp112, Asn134	Asp112
BEL	$\beta$ -galactósidos	$\alpha$	Tyr42	Arg27, Asp44	—
	$\beta$ -galactósidos	$\beta$	Tyr90	Arg74, Asp92	—
	$\beta$ -galactósidos	$\gamma$	Tyr136	Arg121, Asp138	—

\*Solo se han analizado los residuos que interaccionan directamente con el anillo de galactosa.

**Tabla 4.3.** Comparación de los sitios operativos de unión a azúcar entre las lectinas fúngicas con plegamiento tipo trébol  $\beta$  caracterizadas en la actualidad.



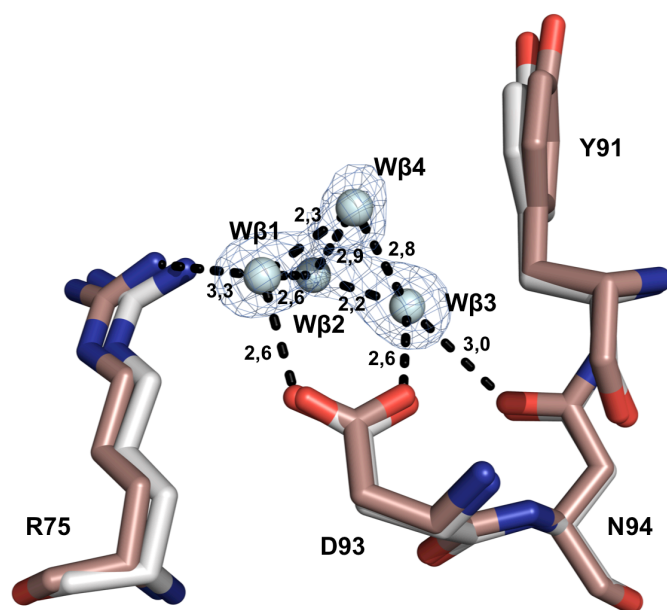
**Figura 4.9.** Vista detallada de los sitios de unión de azúcar en los tréboles  $\beta$  fúngicos CNL (A) y CCL2 (B). En A, la proteína CNL se representa en complejo con lactosa (cian); mientras que en B, la proteína CCL2 aparece en complejo con el trisacárido de quitobiosa fucosilado (GlcNAc $\beta$ (1 $\rightarrow$ 4)[Fuc(1 $\rightarrow$ 3)]GlcNAc) (violeta). Las principales interacciones se representan en línea de puntos con las medidas de las distancias dadas en ångströms (Å).

En la **Tabla 4.3** se muestra el mismo estudio anterior, particularizado para los tréboles  $\beta$  fúngicos cuyas estructuras están resueltas a día de hoy. Si bien se puede ver que el residuo aromático más frecuentemente implicado en la unión de azúcares es la tirosina, también aparecen casos que claramente indican variabilidad en el modo de unión. En ese sentido, cabe destacar a las lectinas CNL y CCL2. Por un lado, la lectina CNL de *C. nebularis* carece del característico residuo aromático y, por tanto, de la interacción de van der Waals característica de esta familia de lectinas. En su lugar, posee una red de enlaces de hidró-

geno que mantienen fija a la molécula de ligando (Pohleven *et al.*, 2012) (**Figura 4.9A**). En cambio, la lectina CCL2 de *C. cinerea* no sigue el modo canónico de unión. En este caso, el ligando descansa en la interfase entre los subdominios  $\beta$  y  $\gamma$ , concretamente, entre las hebras  $\beta_6$  y  $\beta_8$  y el bucle  $\beta_7$ - $\beta_8$  del subdominio  $\beta$ , por un lado, y el bucle  $\beta_9$ - $\beta_{10}$  del subdominio  $\gamma$ , por otro (Schubert *et al.*, 2012) (**Figura 4.9B**).

### 1.3.5. Papel clave de las moléculas de agua en la unión de lactosa

LSL<sub>150</sub> se caracteriza por no presentar cambios conformacionales sustanciales tras la unión del ligando. De hecho, la superposición de las estructuras de LSL<sub>150</sub> y del correspondiente complejo con lactosa proporciona un RMSD para todos los C $\alpha$  de tan solo 0,68 Å. No obstante, el estudio detallado en torno a los sitios de unión de lactosa en ambas estructuras ha demostrado que existen reorganizaciones tanto de las cadenas laterales de algunos aminoácidos como de moléculas de agua implicadas directamente en la unión.



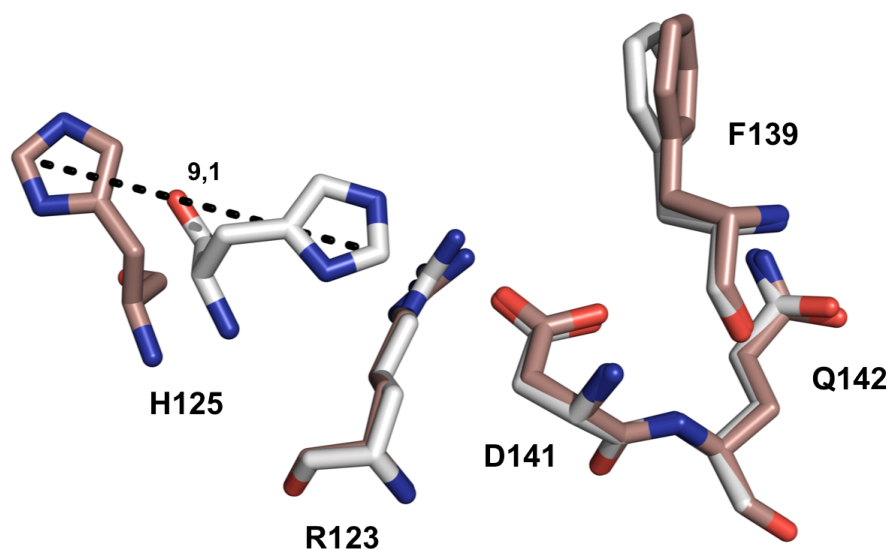
**Figura 4.10.** Comparación estructural del sitio  $\beta$  entre los estados con y sin ligando de LSL<sub>150</sub>. Las cadenas laterales de la estructura de LSL<sub>150</sub> sin ligando vienen representadas en color *marrón*, mientras que las de la estructura con ligando están representadas en *gris*. Las moléculas de agua están representadas mediante *esferas* de color *cian*. El mapa de densidad electrónica  $2F_o - F_c$  (*azul*) a  $1\sigma$  se muestra en torno a las moléculas de agua. Los enlaces de hidrógeno están representados mediante *líneas discontinuas*.

En lo que al sitio  $\beta$  respecta (**Figura 4.10**), los aminoácidos que interaccionan con el anillo de galactosa apenas modifican su posición tras la unión del azúcar. Se puede considerar, por tanto, que en lo que a la proteína respecta, el sitio  $\beta$  está preformado, una situación común en lectinas (Weis y Drickamer, 1996). La estructura de LSL<sub>150</sub> sin ligando revela la existencia de numerosas moléculas de agua ordenadas en este sitio de unión, que se puede afirmar que juegan un importante rol en la unión de los azúcares, creando un campo potencial que favorece la aceptación de las moléculas de ligando (Charcot y Robin, 1853; Lemieux, 1996). Dos de estas moléculas de agua (W $\beta$ 1 y W $\beta$ 3) ocupan posiciones equivalentes a las de los grupos hidroxilo O6 y O4 del anillo de galactosa y poseen unos factores *B*

(8,2 Å<sup>2</sup> para W $\beta$ 1 y 6,2 Å<sup>2</sup> para W $\beta$ 3, respectivamente) similares a los de los átomos de los residuos del sitio de unión, y más bajos que los de las moléculas de agua cercanas (28,7 Å<sup>2</sup> y 24,8 Å<sup>2</sup> para W $\beta$ 2 y W $\beta$ 4, respectivamente) con las que están interaccionando.

Con respecto a las moléculas de solvente del entorno en el que se sitúa el anillo de glucosa, cabe decir que no se puede llevar a cabo un análisis similar al descrito arriba debido a la influencia del empaquetamiento cristalino: en el cristal de LSL<sub>150</sub> sin ligando, este entorno se encuentra parcialmente ocupado por una de las moléculas de proteína relacionada por simetría.

A diferencia de lo que ocurriría en el sitio  $\beta$ , en donde la conformación de los residuos implicados no cambia al unir lactosa, la His-125 del sitio  $\gamma$  experimenta un cambio conformacional significativo por la unión de este azúcar (**Figura 4.11**): la cadena lateral de este residuo se desplaza alrededor de 9 Å y forma enlaces de hidrógeno (átomo Ne-2) con los átomos de oxígeno O3 y O4 del anillo de galactosa. Además, la cadena lateral de la Phe-139 sufre una pequeña reorganización previsiblemente para optimizar el apilamiento con el anillo de galactosa.

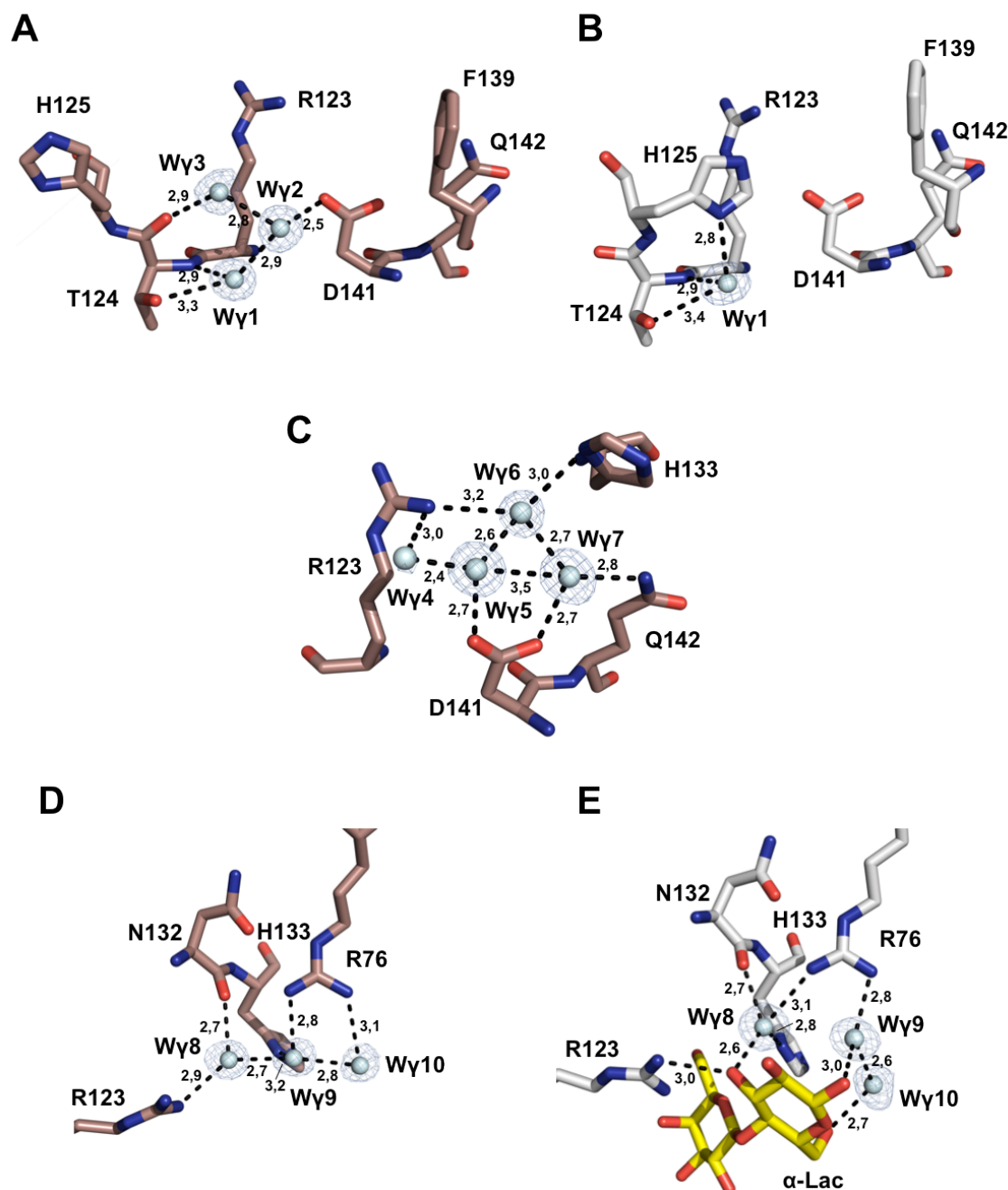


**Figura 4.11.** Comparación estructural del sitio  $\gamma$  entre los estados con y sin ligando de LSL<sub>150</sub>. Las cadenas laterales de la estructura de LSL<sub>150</sub> sin ligando vienen representadas en color *marrón*, mientras que las de la estructura con ligando están representadas en *gris*. El desplazamiento entre las dos posiciones del residuo de His-125 viene indicado con una *línea de puntos* y la distancia viene dada en *ångströms* (Å).

El análisis del solvente en el sitio  $\gamma$  en ambas estructuras ha permitido identificar tres grupos definidos de moléculas de agua (**Figura 4.12**). El grupo 1 (**Figura 4.12A y B**) estaría constituido por las moléculas W $\gamma$ 1, W $\gamma$ 2 y W $\gamma$ 3. Debido a la unión de la lactosa, la His-125 se mueve a su nueva posición, desplazando a las moléculas W $\gamma$ 2 y W $\gamma$ 3, mientras



que W $\gamma$ 1 permanece en la misma posición. Esta molécula de agua queda estabilizada mediante la formación de tres enlaces de hidrógeno con el nitrógeno N $\delta$ -1 de la His-125, el nitrógeno del enlace peptídico y el oxígeno O $\gamma$  de la Thr-124, respectivamente.



**Figura 4.12.** Distribución de las moléculas de agua del sitio  $\gamma$  en los estados con y sin ligando de LSL<sub>150</sub>. El grupo 1 de moléculas de agua se muestra en la estructura sin ligando (**A**) y en presencia de éste (**B**). El grupo 2 se representa en la imagen **C**. Las moléculas de agua del grupo 3 (**D**) experimentan un desplazamiento cuando la molécula de lactosa (*amarillo*) es reconocida por la lectina (**E**). Las cadenas laterales de la estructura de LSL<sub>150</sub> sin ligando vienen representadas en color *marrón*, mientras que las de la estructura con ligando están representadas en *gris*. Las moléculas de agua aparecen como *esferas en cian* y las interacciones se representan en *línea de puntos*, dándose las distancias en *ångströms* (Å). El mapa de densidad electrónica  $2F_0 - F_c$  (*azul*) a  $1\sigma$  se muestra en torno a las moléculas de agua.

Por otra parte, el grupo 2 (**Figura 4.12C**) está formado por cuatro moléculas de agua: W $\gamma$ 4, W $\gamma$ 5, W $\gamma$ 6 y W $\gamma$ 7. Estas moléculas juegan un papel parecido a las identificadas en el sitio  $\beta$ : las moléculas W $\gamma$ 5 y W $\gamma$ 7 ocupan exactamente la misma posición que los grupos hidroxilo O4 y O6 del residuo de galactosa, lo cual se refleja en los valores de los factores  $B$ : 30,7, 6,2, 10,3 y 7,6 Å<sup>2</sup> para W $\gamma$ 4, W $\gamma$ 5, W $\gamma$ 6 y W $\gamma$ 7, respectivamente.

Por último, el grupo 3 (**Figura 4.12D y E**) lo forman las moléculas de agua W $\gamma$ 8, W $\gamma$ 9 y W $\gamma$ 10. Cuando la lactosa se une al sitio  $\gamma$ , las tres moléculas sufren un desplazamiento concertado que las lleva a interactuar con los oxígenos O3', O1' (sólo en el anómero  $\alpha$ ) y O6' del anillo de glucosa, respectivamente. Este comportamiento del solvente es consistente con la idea de que las moléculas de agua se comportan como una extensión de la superficie de la proteína (Toone, 1994). Su comportamiento puede ser descrito como un mecanismo en el que algunas moléculas de agua son sustituidas por grupos hidroxilo del anillo de glucosa para mediar la interacción entre el azúcar y los residuos Arg-132, His-133 y Arg-76. Aunque este mecanismo puede ser explicado también por una liberación de las moléculas de agua tras la unión de la lactosa y la posterior unión de nuevas moléculas de solvente, creemos que el mecanismo propuesto es más simple porque implica menores cambios en la reorganización de las moléculas de agua y de los enlaces de hidrógeno ya existentes.

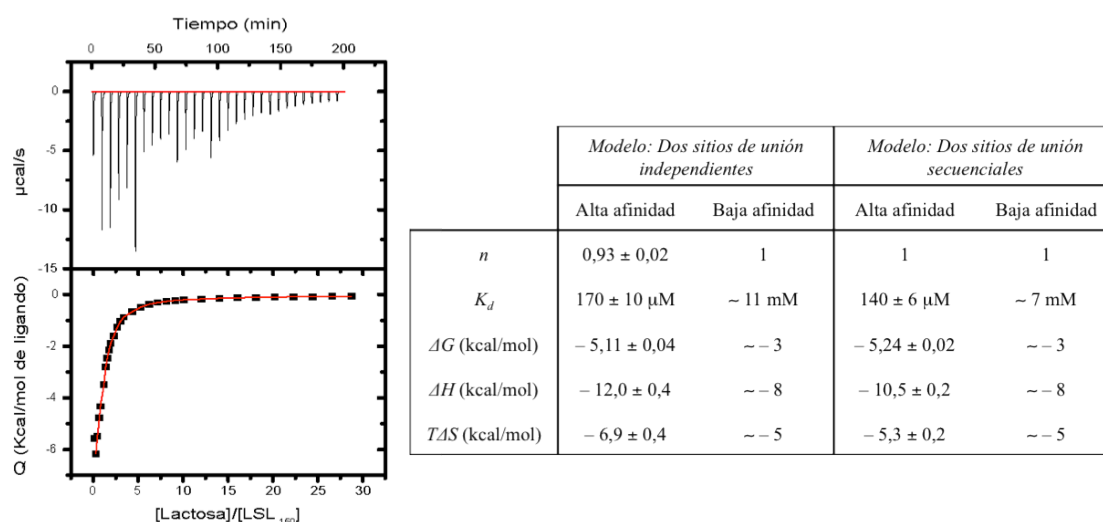
El estudio detallado de la distribución de las moléculas de agua en los dos estados de LSL<sub>150</sub>, con lactosa y sin ella, se ha podido llevar a cabo gracias a la resolución final obtenida en ambos casos (1,67 y 1,47 Å, respectivamente). De hecho, los grupos de moléculas de agua considerados más arriba están perfectamente definidos en el mapa de densidad electrónica experimental. Bien es cierto, no obstante, que sólo podemos hacer una descripción parcial del fenómeno de la unión del ligando y de la reorganización experimentada por las moléculas de agua, puesto que cristalográficamente únicamente observamos aquellas moléculas de solvente que se encuentran más ordenadas en el cristal (Angulo *et al.*, 2011).

Aunque la cristalografía de rayos X brinda la posibilidad de observar las posiciones que ocupan las moléculas de solvente en la superficie de las proteínas e identificar algunas de las interacciones en las que participan, no deja de ser una interpretación parcial del esquema completo. La contribución termodinámica (Gabiús, 2009), las propiedades electrostáticas que otorgan a la superficie de la proteína (Nadig *et al.*, 1998) y los tiempos de residencia en las diferentes posiciones son otros aspectos que hay que considerar para completar el estudio desde otras aproximaciones.



### 1.4. La afinidad de LSL<sub>150</sub> por lactosa en solución es baja

La capacidad del módulo lectina LSL<sub>150</sub> para unir lactosa en solución fue investigada mediante calorimetría isotérmica de titulación (ITC). Los experimentos de ITC se llevaron a cabo a 25 °C en tampón TSA utilizando lactosa a una concentración final de 30 mM. El perfil de unión que se obtuvo experimentalmente (**Figura 4.13**) demuestra que la unión a lactosa es exotérmica, como indica la orientación de los picos en la gráfica  $\mu\text{cal/s}$  versus tiempo. A la vista de la curva de titulación, se observa cómo ésta se desvía claramente del comportamiento esperado para un modelo de un solo tipo de sitios de unión, lo que sugiere, por tanto, la presencia de, al menos, dos tipos diferentes de sitios.



**Figura 4.13.** Análisis mediante ITC de la unión en solución de lactosa a LSL<sub>150</sub>. A la izquierda se muestra la curva de valoración, así como el ajuste de los datos (rojo) con el modelo de dos sitios de unión independientes. A la derecha se muestra una tabla con los parámetros termodinámicos calculados a partir de la aplicación de dos modelos de ajuste distintos: (i) dos sitios de unión independientes y (ii) dos sitios de unión secuenciales. Los parámetros termodinámicos determinados han sido: la estequiometría ( $n$ ), la constante de disociación ( $K_d$ ), la energía de Gibbs ( $\Delta G$ ), la entalpía de unión ( $\Delta H$ ) y el cambio de entropía ( $T\Delta S$ ).

El análisis de la isoterma de unión con un modelo que considera dos tipos de sitios de unión independientes indica que los dos sitios de unión de LSL<sub>150</sub> mostrarían afinidades muy distintas por lactosa. De este modo, habría un sitio de alta afinidad relativa con una constante de disociación ( $K_{d1}$ ) de  $170 \pm 10 \mu\text{M}$  ( $\Delta G_1 = -5,11 \pm 0,04 \text{ kcal/mol}$ ) y una entalpía ( $\Delta H_1$ ) de  $-12,0 \pm 0,4 \text{ kcal/mol}$ , cuya estequiometría es prácticamente 1 ( $0,93 \pm 0,02$ ) y, por otro lado, habría un sitio de baja afinidad relativa para el cual sólo se pudo hacer una estimación aproximada de los valores de los parámetros termodinámicos debido a que, en las condiciones del ensayo, no se consiguió la saturación completa del sistema. Ajustando manualmente la estequiometría de este segundo sitio a 1, se obtuvo una constante de disocia-

ción ( $K_{d2}$ ) de  $\sim 11$  mM ( $\Delta G_2 = \sim -3$  kcal/mol) y una entalpía ( $\Delta H_2$ ) de  $\sim -8$  kcal/mol. En ambos sitios se observa que la unión del disacárido está dirigida entálpicamente, existiendo una penalización entrópica:  $T\Delta S_1 = -6,9 \pm 0,4$  kcal/mol y  $T\Delta S_2 = \sim -5$  kcal/mol. Este comportamiento es típico de las interacciones lectina-carbohidrato (Gabijs, 2009). Las variaciones de entalpía estarían justificadas por la formación neta de interacciones favorables entre el ligando y la proteína. Estas interacciones (enlaces de hidrógeno e interacciones apolares) compensarían de alguna manera la pérdida de libertad translacional y conformacional que sufren tanto el ligando como las cadenas laterales de los aminoácidos implicados en el reconocimiento del azúcar (Dam y Brewer, 2002). Además de la contribución de las interacciones entre la proteína y el ligando a la variación de entalpía, se ha demostrado que la liberación de moléculas de agua al solvente también es un fenómeno favorable no sólo entrópicamente. Así, la reducción de la superficie expuesta al solvente experimentada por la unión del ligando y la consecuente liberación de moléculas de agua, como había sido descrito en el apartado anterior, contribuye favorablemente a la entalpía de unión por la formación de enlaces de hidrógeno más favorables entre moléculas de solvente (Toone, 1994).

Aplicando un modelo de interacción alternativo, que considera la presencia de dos sitios de unión secuenciales, se obtiene también un buen ajuste de los datos, similar al obtenido con el anterior modelo, con valores de los parámetros termodinámicos muy similares (**Figura 4.13**). Sin embargo, teniendo en cuenta que el anterior modelo implica menos condicionantes y que la información estructural que poseemos lo sustenta, consideramos no justificado el uso de un modelo secuencial. Además, cuando las afinidades de ambos sitios de unión en LSL<sub>150</sub> distan notablemente entre sí, se puede considerar que no existe una diferencia apreciable entre ambos modelos.

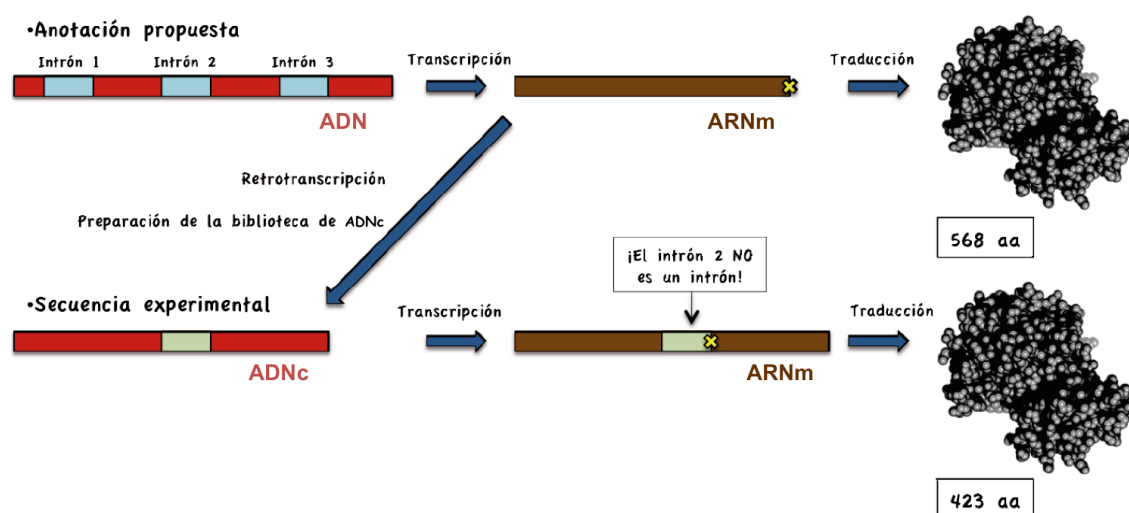
Como hemos indicado, los resultados de ITC son consistentes con la información estructural obtenida mediante cristalografía de rayos X. La presencia de dos sitios de unión con distintas afinidades coincidiría con la existencia de los dos sitios de unión en la estructura de LSL<sub>150</sub>. La diferente afinidad debería explicarse, por tanto, por las distintas interacciones observadas en ambos sitios. Teniendo en cuenta que en el sitio  $\gamma$  aparece un mayor número de interacciones, éste podría ser el sitio con mayor afinidad, en tanto que el sitio  $\beta$  sería el de menor afinidad. En este contexto, por ejemplo, el sitio  $\gamma$  posee los residuos de His-125 e His-133 que actúan como unas pinzas, ayudando a establecer mayor número de contactos que los observados en el sitio  $\beta$  (**Figura 4.8**). Además, esta asignación está sustentada por resultados previos de las estructuras de los complejos de la lectina completa

LSLa en las que sólo el sitio  $\gamma$  estaba ocupado por moléculas de lactosa o N-acetilactosamina (LacNAc). De hecho, la molécula de LacNAc en el sitio  $\gamma$  se observaba cuando se llevaba una inmersión rápida de los cristales de LSLa con una solución crioprotectora que contenía una concentración 20 mM de ligando. En cambio, para observar una molécula de LacNAc en el sitio  $\beta$ , se necesitaba incubar los cristales durante toda la noche con una concentración mucho más elevada (0,1 M) (Mancheño *et al.*, 2005).

## 2. La lectina LBL de *Laccaria bicolor*

### 2.1. El gen *lbl* codifica una proteína distinta a la predicha

El genoma del hongo *Laccaria bicolor* (Martin *et al.*, 2008) contiene una fase de lectura abierta (ORF, *open-reading frame*) que codificaría una proteína (LBL) con un dominio lectina en su extremo N-terminal que poseería un porcentaje de identidad de secuencia del 52% con el módulo lectina N-terminal de LSLa. Este gen, denominado *lbl*, fue amplificado a partir de una biblioteca de ADNc, proporcionada por el Dr. Francis Martin (apartado 2.3 de *Materiales y Métodos*). Además de este gen, también se amplificó el fragmento, que denominamos *lbl*<sub>152</sub>, que codificaría el módulo lectina de dicha proteína y que se corresponde con sus 152 primeros aminoácidos (LBL<sub>152</sub>). Los dos fragmentos amplificados fueron clonados, obteniéndose así los plásmidos pKLBL y pKLBL<sub>152</sub>, respectivamente.



**Figura 4.14.** Esquema que muestra la repercusión directa sobre la secuencia de aminoácidos de la lectina de *L. bicolor* al no procesarse el intrón 2, tal y como ocurre en la secuencia experimental.

ATGTCAAACGAGTACAATCCCCCTCTTGAATTCGCTTTTCGTCTGTGCGGCCCTTGCAGCGACCGCGTCCTTTTTTCTCGAGTTTC  
 CCCATCTCCGGAAGTGTTCACCATCCCAAAAGTGAAAGTTTATCCCGACAGTGGTTTGTTCGCCATTCAGGCTCGGGCCAAAACG  
 CAGGATGTTACGCCATTAAAAAGCAAGAACTGGGAAGGTCTCTTTTCCAGAAATGTGCGCTGATCCTCGCGTTGGACACATTGAC  
 GCGACGGTAAATACCCAGACAATTGGTTCAAGTTTGAAGCAGGCAAGTACGAGGCTACTTCCGCTTGGCGCGCGTCCG  
 ATCAGACACTGTCCTGGTGTCTAGAACTTCGACGGGCACTGACACCCAGGTTATCAACTACCCGGCCACCAGTGCAAAATACGATG  
 ATCAATATTTCACTATCTTATTCGACAAGGACCAGCTATCTACGAAAGGCCAGTCAACTTGGGCGGCTCCAGTATGCTCGTCCT  
 CGCATTACCACCTCTTGACGCATCTCGACTCCCCACTCCCCCGGAGGCTGGAAATTCGCCAACCTTGTTAATCCAGCAAGCCTGA  
 TTCTAACCTTTTCACTACCCAGTCTTACTCTTTGGCAGCTACACATACACCG**GTGCGCGATCTTCTCTCATTGTACATTTTAT**  
**TCTGACTTGTGAATCCGCATACAG**CTCTCAGCAGCGCCAAGGACGACAATCGCATGAAGATCGTTGCTTGGGATAGGAACCGAAC  
 CTTGTCAAGGAGCTGGAAAAGCCAGGTGCCCGGTACCTTTGGAATATTACGTTCAACGCTGGAAACCAGACGGTGACCTTCACCGG  
 GCAATCTAATCAATCCATTACAGGCCACTCTAGATGAAGTGGTTGTCAGGCGCCGCGTGGTTTACCTTCCATCGTCTCAGGTCCCA  
 TCCTTCCGTACACTCTCACCTACGCCGCGGCTGACCCATCTCGCTTCCCTGTCAATTAAGAACGGCCCTTACTCCTTCTGGCCTGCG  
 TCATTCAAGGACGGGCGTAAATCCTTCTGTATTGCCGTCGCCGACAGAGATAATCTGGTTCAGAAAGTTGATTGAGTGCCGAGGCGG  
 AAAGGTCAATTCGAGATGTCCAGGTCAATGACGCTCAACGCACGATTGGGCTCATTTGGCCAGGATGGTTCCAAGGCTACTTTCAACT  
 ACGACATTGCCCTAACATGCTTCTACGACTTCT**GCAACTTCACCAAGGATGCCTTCCTCCTTCTCGCAAAGGATTTTCGATGTAAC**  
**CTTTCGGATACTAACGCTGCATTCCGGGTTCCTCATGAAG**GGCTCCTGGTGTGTTGGCCCTGGTCTCATCGCCGGCGCTGCGGCAT  
 CGGATTGGATCGACCGTACATTCACGGTATAAGCTCTCAGCTTTGGAGAGGGCTTCTCTTGTAGG**GTAAAGCGATCATAGTGCAA**  
**TATCATGTGGATCTTCGCTTACGCATTTTTT**AGCTACTCCAAGTTCCTTGGGCGCTTTGACAACACAGGAGGAGGAGGATA  
 TCTTCGATTCTCCCTTCCCTCGCTACCAAGAAGTACCGTATCTGCTGCCGTTGTTGAATCTGGGTGGTCCCGTATCGTCGACAGC  
 ACGCTCATCAGCACCAAGAACCCGTCATCTCGCGAAACACATTCTTCGTGACCTCGTCTCGCATCTCAACCAATACCTGCGACAA  
 TCGTCAGTCATGCCATACCCCTTGCAACTTCCAACGCACCTCATACTTTGCCCGCTCACCATCCGCCCTTTTGAGATTCTCAGCT  
 GCGCTCCTCATCAGCTCGACCAACACATCCCGGGCAAAGTCTTCTGGACTTTGTGTCTGTTCTCGTCTTTCATGAGATCGGACCGC  
 TGTTATTGGTGTCTTTAA

## B

LBL teór: 1 MSNEYNPPLGIAFRLCGLASDRVLSRVSPSEVFHHPKSEVYPDQWFVAIPGSGQNAGC 60  
 LBL exp: 1 MSNEYNPPLGIAFRLCGLASDRVLSRVSPSEVFHHPKSEVYPDQWFVAIPGSGQNAGC 60

LBL teór: 61 YAIKSKNTGKVLFSRMSDPDRVGHIDGDGKYPDNWFKFEAGSGKYAGYFRLRAVASDTVL 120  
 LBL exp: 61 YAIKSKNTGKVLFSRMSDPDRVGHIDGDGKYPDNWFKFEAGSGKYAGYFRLRAVASDTVL 120

LBL teór: 121 VSRTSTGTDQVINYPATSAKYDDQYFTILFDKDQLSTKGQSNLGGLOQYARPRITTLDS 180  
 LBL exp: 121 VSRTSTGTDQVINYPATSAKYDDQYFTILFDKDQLSTKGQSNLGGLOQYARPRITTLDS 180

LBL teór: 181 RLPTPPGGWKFANLVNPSKPDNSLHYPVLLFGSYTYTALSSAKDDNRMKIVAWDRNRNL 240  
 LBL exp: 181 RLPTPPGGWKFANLVNPSKPDNSLHYPVLLFGSYTYTALSSAKDDNRMKIVAWDRNRNL 240

LBL teór: 241 VKELEKPGARYLWNITFNAGNQTVTFTGQSNQSIQATLDELVVRPAVVYLPSSSGPILPY 300  
 LBL exp: 241 VKELEKPGARYLWNITFNAGNQTVTFTGQSNQSIQATLDELVVRPAVVYLPSSSGPILPY 300

LBL teór: 301 TLTYAAADPSRFPVIKNGPYFSPASFQDGRKSFCIAVADRDNLVQKLECPGGKVIRDV 360  
 LBL exp: 301 TLTYAAADPSRFPVIKNGPYFSPASFQDGRKSFCIAVADRDNLVQKLECPGGKVIRDV 360

LBL teór: 361 QVNDQRTIGLIGQDGSKATFNVDIALTCFYDFWLLVVGPGLIAGAAASGLDRPYIPRYK 420  
 LBL exp: 361 QVNDQRTIGLIGQDGSKATFNVDIALTCFYDFCNFTKDAFLLLAKDFDVTLSDTNAAFR 420

LBL teór: 421 LSALERASLVGYSKFPDGPFDNTGRREDIFDSPFPYQELPYSAAVVESGWSRIVDSTLI 480  
 LBL exp: 421 VPH----- 423

LBL teór: 481 STKNPSISRNTFFVTSSRISTNTCDNRQSWPYPLQLPHTLILCPLTIRPFEILSCAPHQL 540  
 LBL exp: -----

LBL teór: 541 DQHIPGKVFWTLCLFSFFMRSDRCYWCL 568  
 LBL exp: -----

**Figura 4.15. A.** Secuencia del gen *lbl* sin procesar. En **amarillo** aparecen las secuencias de los intrones; en **rojo**, los pares de bases del comienzo y del final de los intrones, respectivamente; y en **verde**, las *branch sequences*. En el supuesto intrón 2 aparece en **azul** el codón de terminación experimental del gen que codifica la lectina de *L. bicolor*, determinado por nuestro grupo. **B.** Alineamiento entre la secuencia teórica y experimental de la lectina de *L. bicolor*. En color **azul** se encuentra el dominio lectina N-terminal propuesto; y sombreado en **amarillo**, los diferentes extremos C-terminal que aparecen según el procesamiento o no del intrón número 2.

Los plásmidos pKLBL y pKLBL<sub>152</sub> se secuenciaron para confirmar ambas secuencias. Sorprendentemente, mientras que la secuencia de *lbl*<sub>152</sub> correspondía con lo esperado según la secuencia genómica, no ocurrió lo mismo con la secuencia de *lbl*. Al hacer un alineamiento de secuencias entre el fragmento insertado en la construcción pKLBL y el gen que

se encontraba en la base de datos (LACBIDRAFT\_318163), se observó que la secuencia experimental presentaba un segmento adicional de cerca de 94 pb próximo al extremo 3'. Considerando esta discrepancia, que podría indicar una predicción incorrecta de los posibles intrones presentes en el gen que codifica la lectina en el genoma de *L. bicolor*, se procedió a comparar la secuencia experimental (secuencia procesada) con la del genoma de *L. bicolor* (secuencia sin procesar y, por lo tanto, conteniendo intrones) (**Figuras 4.14 y 4.15**). Concretamente, el análisis del fragmento o cóntigo número 610 (LACBIscfold\_8\_Cont610) ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) permitió identificar cuál es la secuencia original del gen de la lectina. Los resultados de la comparación de ambas secuencias arrojaron como resultado que uno de los tres intrones no era tal, apareciendo en el ARN mensajero maduro (ARNm). En efecto, el supuesto intrón 2 en realidad permanece en la secuencia de ARNm, lo que hace que aparezca un codón de terminación antes del propuesto en la base de datos y que, por tanto, la secuencia de aminoácidos de la lectina LBL de *L. bicolor* sea distinta a la predicha. La secuencia experimental codificaría una proteína de 423 residuos (masa molecular teórica de 46,8 kDa). La posición de este hipotético intrón no afecta a la secuencia del fragmento N-terminal *lbl*<sub>152</sub>.

Aunque inicialmente el intrón número 2 se clasificase como tal, su secuencia presenta diferencias con las de los intrones 1 y 3 (**Figura 4.15**). Mientras que las secuencias de los intrones 1 y 3 comienzan con el par GT y acaban con el AG, la del 2 empieza por GC y acaba en AG. Se sabe que las secuencias consenso en los extremos 5' y 3' de los intrones son críticas para el correcto procesamiento, ya que el cambio en una única base podría inhibir el procesamiento (*splicing*) durante la maduración del ARNm (Clancy, 2008). Además, la denominada *branch sequence*, clave en el mecanismo de escisión del intrón, aparece cerca del extremo 3' en la secuencia 2 y no en el centro de la secuencia, como ocurre en la mayoría de intrones (Konarska *et al.*, 1985).

La correcta identificación de los genes en un genoma se basa principalmente en estudios comparativos y de similitud de secuencia y es conocido que esta aproximación y la subsiguiente anotación presenta importantes limitaciones (Valencia, 2005; Brent, 2008). Las modificaciones postranscripcionales, la compartimentalización celular, así como otros mecanismos de regulación influyen en que cerca de un 50-70% de los ORF identificados en genomas secuenciados no tengan una función asignada o estén caracterizados solo superficialmente (Raes *et al.*, 2007). Por esta razón, la identificación correcta de la distribución de exones e intrones en la estructura génica es imprescindible para llevar a cabo interpretaciones correctas sobre la estructura y función de las proteínas. Este proceso es particular-

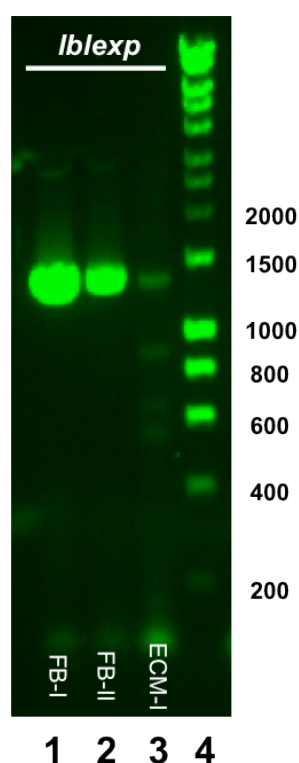
mente complejo en el caso de los genomas de hongos basidiomicetos debido a la elevada densidad de intrones y de exones que codifican secuencias incluso de menos de diez aminoácidos (Kupfer *et al.*, 2004; Stajich *et al.*, 2007). Como cabría esperar, al ser *L. bicolor* un hongo basidiomiceto, su genoma presenta el comportamiento típico de este grupo. De hecho, un estudio llevado a cabo por Larsen y colaboradores, basado en la aplicación de métodos bioinformáticos que estudian detalladamente las secuencias de ARN, indica que el 69% de un conjunto de genes involucrados en la micorrización en *L. bicolor* se encontraba mal anotado (Larsen *et al.*, 2010). Aun así, habiendo servidores de estimación de intrones [por ejemplo, Augustus (Stanke y Morgenstern, 2005)], la inspección manual y experimental confirmarán siempre en último lugar la secuencia correcta para un gen determinado.

Antes de llevar a cabo la amplificación del nuevo gen *lblexp* (secuencia experimental del gen *lbl*), se verificó que el intrón número 2 se mantenía en el ARNm maduro, independientemente de la biblioteca de ADNc de la que se partiera: FB-I (estadio temprano) o FB-II (estadio tardío). El resultado fue el mismo en los dos casos, por lo que quedaba claro que el procesamiento del ARNm de LBL no depende de la etapa de maduración del hongo. Conocida la correcta secuencia del gen *lbl*, éste se volvió a amplificar con un nuevo par de oligonucleótidos a partir de la biblioteca de ADNc para su posterior clonación, obteniéndose el plásmido pKLBLExp, que se utilizó de aquí en adelante. Los resultados obtenidos fueron comunicados al Dr. Francis Martin, principal responsable de la anotación del genoma de *L. bicolor*, quedando anotada la modificación de la secuencia en lo que al procesamiento del segundo intrón respecta.

## 2.2. La lectina LBL se hiperproduce en los cuerpos fructíferos

*L. bicolor* es un hongo basidiomiceto y ectomicorrícico, lo que significa que establece relaciones simbióticas con las raíces de muchas especies arbóreas, como álamos y abetos (Villeneuve *et al.*, 1991). Hasta ahora muchos de los genes identificados en el genoma de *L. bicolor* se encuentran asociados con los procesos de simbiosis y de alimentación saprófita como, por ejemplo, proteasas, transportadores e hidrolasas extracelulares específicas de polisacáridos (Martin y Nehls, 2009). Generalmente, la localización subcelular de las proteínas en el hongo puede dar pistas acerca de la función que desempeñan y del proceso biológico en el que están involucradas. Por esta razón, a la hora de dilucidar la función es interesante identificar el patrón de expresión de una proteína nueva en los diferentes tejidos del hongo.

Para analizar los niveles de expresión cualitativamente, se amplificó mediante PCR el gen *lblexp* (1.272 pb) a partir de tres bibliotecas de ADNc diferentes: FB-I, FB-II y ECM-I (apartado 2.3 de *Materiales y Métodos*). Los resultados de la PCR se analizaron mediante electroforesis (Figura 4.16). Los resultados indican que la lectina se hiperproduce en los cuerpos fructíferos o basidiocarpos, mientras que los niveles de expresión son mucho más bajos en la ectomicorriza. Además, la producción es mayor en el estadio temprano, lo que indicaría que en principio los niveles de expresión son dependientes del estado de maduración del hongo.



**Figura 4.16.** Análisis mediante electroforesis de agarosa al 1% de los niveles de expresión del gen *lblexp* del hongo *Laccaria bicolor*. El gen ha sido amplificado a partir de tres bibliotecas de ADNc distintas: FB-I (cuerpos fructíferos en un estadio temprano), FB-II (cuerpos fructíferos en un estadio tardío) y ECM-I (ectomicorriza de tres meses entre el hongo y un álamo).

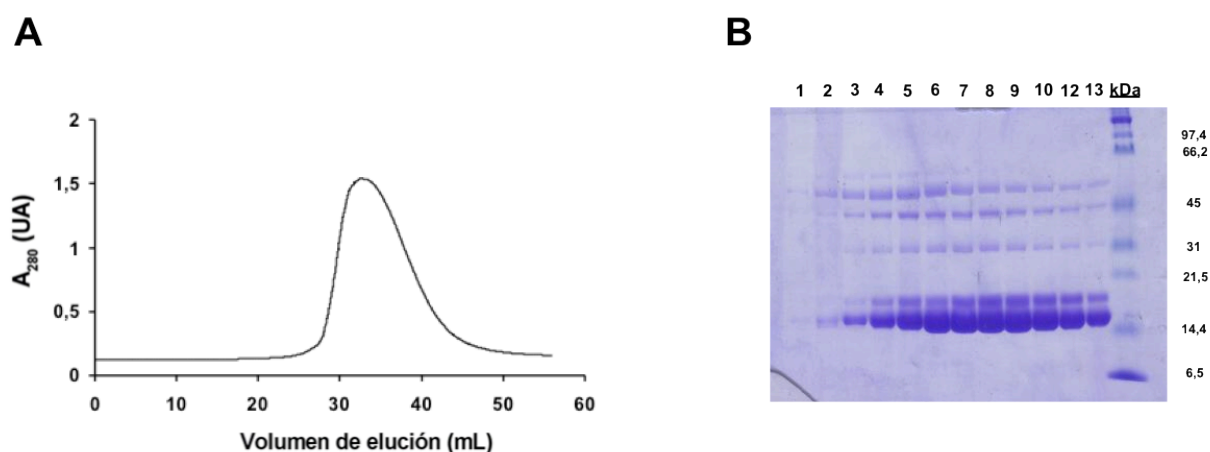
De este modo, probablemente la lectina no está implicada en los procesos de simbiosis del hongo con otras especies. Ciertamente, este comportamiento está de acuerdo con lo observado en otras lectinas fúngicas con estructura trébol  $\beta$ , las cuales se aíslan preferentemente de los cuerpos fructíferos, pues es donde se muestran los niveles de expresión más elevados. En este sentido, recientemente, se ha caracterizado un nuevo grupo de lectinas fúngicas, conocidas comúnmente como lectinas de los cuerpos fructíferos o FBL (del inglés *fruiting body lectins*), que muestran un papel importante en los mecanismos de defensa de los hongos multicelulares frente a parásitos y depredadores debido a su toxicidad (Trigueros *et al.*, 2003; Goldstein y Winter, 2007; Wohlschlager *et al.*, 2011; Schubert *et al.*, 2012). La lectina de *L. bicolor* podría pertenecer, por tanto, a este grupo de proteínas. No obstante, la dependencia con el estadio de crecimiento del hongo podría igualmente indicar

que la lectina estuviese involucrada en algún proceso relacionado con el desarrollo del basidiocarpo.

## 2.3. LBL se proteoliza cuando se produce de forma recombinante

### 2.3.1. Producción de la lectina de *L. bicolor* en células de *E. coli* BL21(DE3)

La purificación de la lectina LBL de *L. bicolor* se ha llevado a cabo aplicando el mismo protocolo que el utilizado con LSL<sub>150</sub>. Como se puede apreciar en la **Figura 4.17A**, el perfil de elución muestra un pico ancho y prácticamente simétrico. El análisis mediante SDS-PAGE de las fracciones (**Figura 4.17B**) revela una complejidad mayor de la esperada. Haciendo un análisis detallado de las bandas que aparecen en los diferentes carriles, se observa que hay fundamentalmente cinco bandas. La banda mayoritaria correspondería con una masa molecular entre 14,4 y 21 kDa, similar a la del módulo lectina N-terminal (16,8 kDa), lo que indicaría que LBL se ha proteolizado. La banda más intensa de la fracción soluble del extracto celular (dato no mostrado) también migra a la misma altura, por lo que el procesamiento proteolítico probablemente esté teniendo lugar antes de la lisis celular, ya que teóricamente el cóctel de inhibidores de proteasas impediría que la proteólisis fuese tan considerable.



**Figura 4.17.** Purificación de la lectina LBL. **A.** Perfil de elución de LBL de la cromatografía de afinidad con Sepharose® 4B. **B.** SDS-PAGE (12% de poliacrilamida) de la purificación de LBL eluida con 0,2 M lactosa de la matriz cromatográfica. Los carriles 1-13 se corresponden con el pico observado en el perfil de elución.



Otra de las bandas que más intensidad muestra es la que migra a la altura de los 45 kDa. Teniendo en cuenta que LBL tiene una masa molecular de 46,8 kDa, esta banda podría corresponderse con la proteína completa. El resto de bandas podrían corresponder a versiones de LBL parcialmente hidrolizadas. La proteólisis en las lectinas, tanto obtenidas de la fuente original como de forma recombinante, es un fenómeno bastante común (Harada *et al.*, 1990; Rodríguez-Arango *et al.*, 1992). Por ejemplo, se ha visto que muchas lectinas recombinantes de plantas, al ser sobreexpresadas en *E. coli*, sufren un procesamiento del extremo C-terminal por distintos puntos que, además, no suelen coincidir con el punto de corte que presentan cuando éstas se extraen directamente de la fuente natural (Young *et al.*, 1995). En este sentido es interesante destacar el paralelismo en la producción recombinante de LBL y LSLa (Tateno y Goldstein, 2003). En ambos casos, las lectinas se proteolizan por sus extremos C-terminales, generándose especies que conservan la capacidad para interaccionar con la matriz de Sepharose® 4B y, por tanto, conteniendo el módulo lectina N-terminal.

Para confirmar que las diferentes bandas obtenidas en la purificación de LBL son fragmentos proteolizados de la misma, las bandas del gel fueron recortadas y analizadas por espectrometría de masas para así obtener las huellas de masa peptídica y de fragmentación peptídica. La combinación de los resultados proporcionados por ambos abordajes posibilita la identificación de proteínas, aunque éstas muestren un procesamiento complejo. Los resultados obtenidos indicaron que todas las bandas que aparecen por debajo de la banda de los 45 kDa en el gel se corresponden con fragmentos que contienen el módulo lectina completo de la proteína LBL, puesto que los péptidos trípticos identificados coinciden con los teóricos derivados de la lectina de *L. bicolor*.

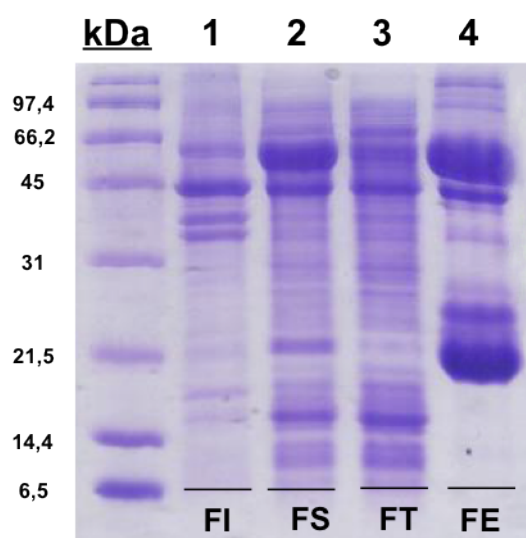
La única banda que aparece por encima de los 45 kDa también fue analizada mediante el análisis proteómico. El resultado identificó a esta banda como GroEL de 55,2 kDa de masa molecular. La masa de la proteína encajaría perfectamente con la migración observada de la banda en el gel (**Figura 4.17B**). GroEL es una chaperona que aparece, tanto en *E. coli* como en otras bacterias (Zeilstra-Ryalls *et al.*, 1991), formando un complejo cilíndrico activo de dos anillos que contienen siete subunidades de GroEL cada uno. Junto con la proteína GroES, GroEL forma un complejo de aproximadamente 1 MDa denominado GroEL/GroES (Xu *et al.*, 1997). Este complejo tiene por función ayudar a que otras proteínas se plieguen adecuadamente, sobre todo cuando éstas exponen residuos de carácter muy hidrofóbico al medio (Mayhew *et al.*, 1996). La actividad de GroEL/GroES es depen-

diente de ATP y generalmente sólo realiza su función sobre proteínas que no superan los 70 kDa por término medio.

Estos resultados apuntan a que LBL necesita de la asistencia del sistema GroEL/GroES para plegarse adecuadamente en *E. coli* (Structural Genomics *et al.*, 2008), pero también que es la parte C-terminal de la lectina la que interacciona con el complejo de chaperonas y que previsiblemente se encuentra plegada incorrectamente. Quizás, en nuestras condiciones experimentales la cantidad de GroEL/GroES es limitante para el correcto plegamiento de LBL, generándose especies parcialmente plegadas, sensibles a la acción de proteasas y que serían las responsables del patrón electroforético observado.

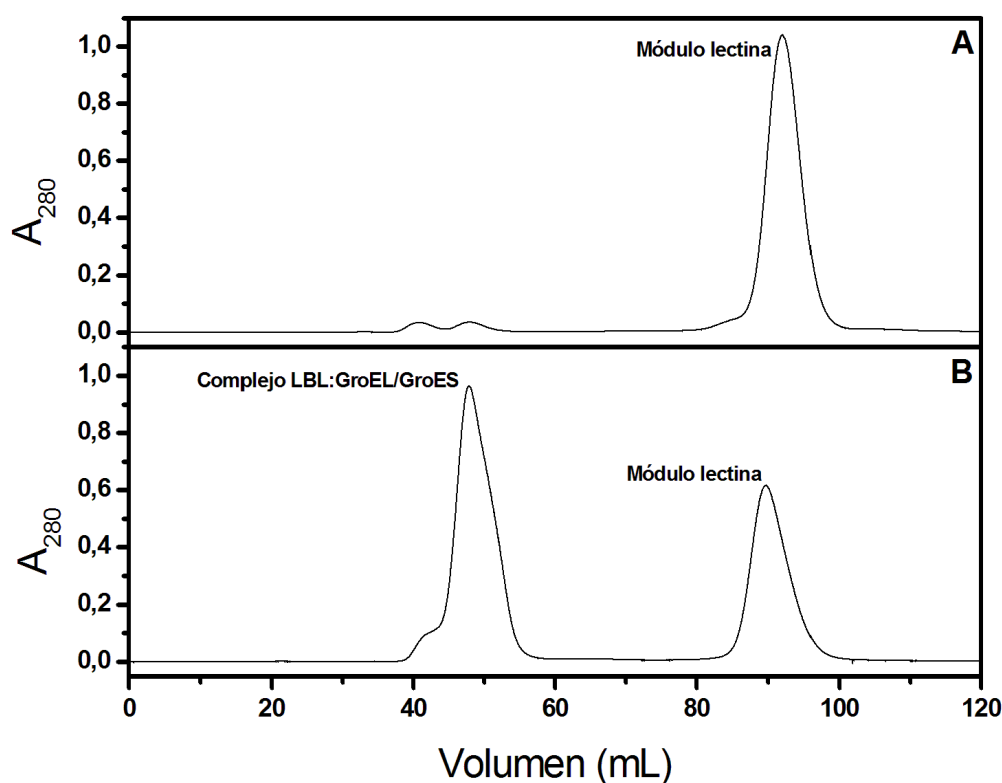
### 2.3.2. Producción de la lectina de *L. bicolor* en células de *E. coli* BL21(DE3) que sobreexpresan el complejo GroEL/GroES

La cotransformación de células de *E. coli* BL21(DE3) con los vectores pKLBLExp y pGro7 nos permitió obtener una cepa con capacidad para sobreexpresar tanto la lectina LBL como el complejo GroEL/GroES (ver sección 2.3.1 de *Materiales y Métodos*). El análisis electroforético de la purificación de la lectina (**Figura 4.18**) reveló que ambas proteínas se expresaban con elevados rendimientos. En estas condiciones, previsiblemente la producción de la chaperona dejaría de ser un factor limitante para el correcto plegamiento del extremo C-terminal de la lectina. No obstante, aun en estas condiciones experimentales, se sigue detectando el complejo entre LBL y GroEL/GroES, pues ambas proteínas coeluyen de la Sepharose® 4B (*carril 4* de la **Figura 4.18**), lo cual muestra el plegamiento incorrecto del extremo C-terminal de LBL, explicando de nuevo la proteólisis de LBL observada.

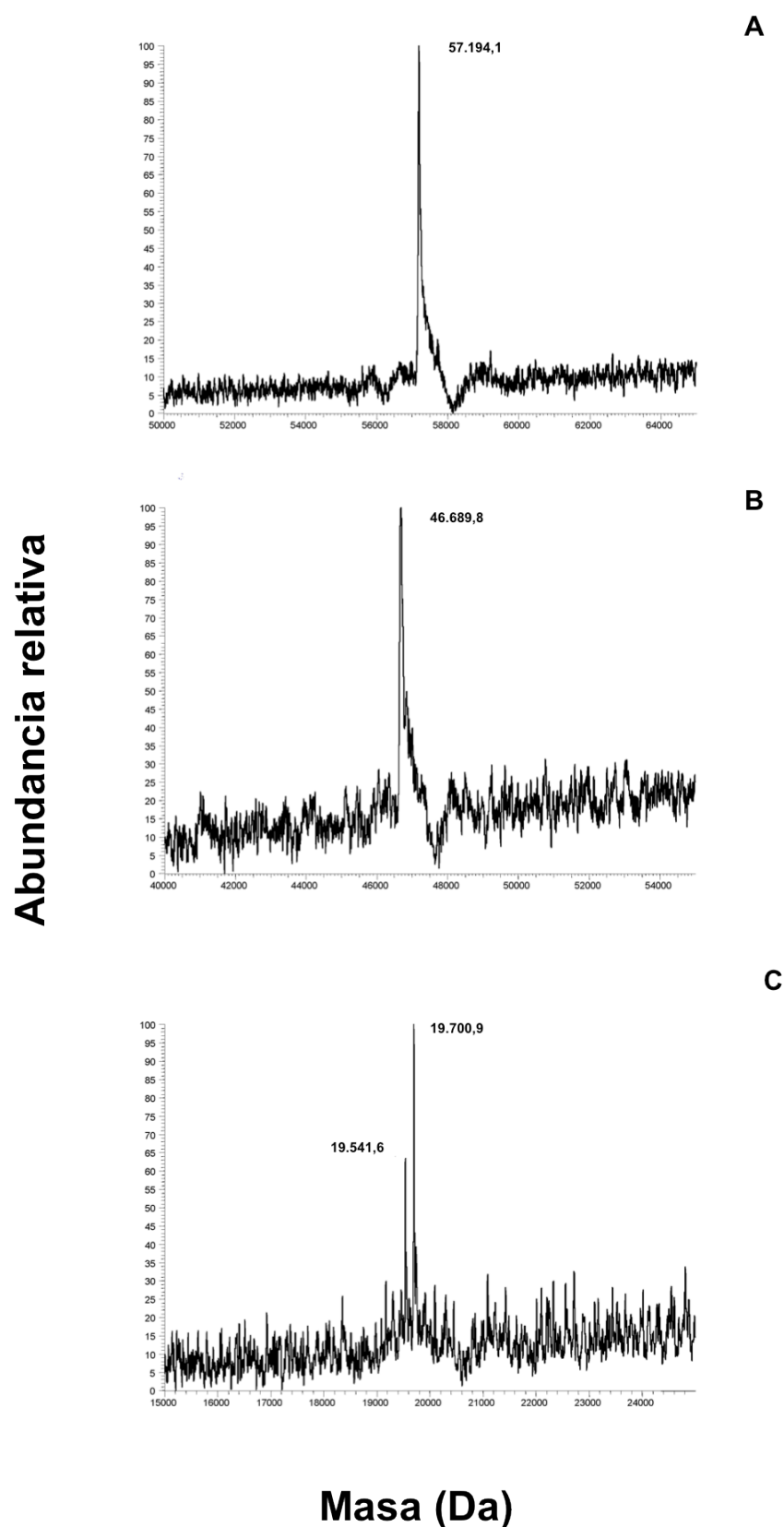


**Figura 4.18.** Análisis electroforético (SDS-PAGE al 12% de poliacrilamida) de la expresión y purificación de LBL en presencia del complejo GroEL/GroES. El carril 1 se corresponde con la fracción insoluble del lisado celular (FI), el carril 2 con la fracción soluble del extracto celular (FS), el carril 3 con la fracción no retenida a la columna o *flow-through* (FT), y el carril 4 con la fracción purificada al lavar con el tampón de elución TSAL (FE).

Las fracciones obtenidas en este primer paso de purificación cromatográfico se concentraron y analizaron mediante cromatografía de exclusión molecular, usando una columna Superdex 200 PG (GE Healthcare Life Sciences). El perfil de elución obtenido se comparó con un control en el que se había producido LBL como se ha descrito previamente en la sección 2.3.1 de *Resultados y Discusión* (**Figura 4.19**). El resultado de la comparación indica que existen al menos dos especies en solución: una que se corresponde con el módulo lectina N-terminal (pico de la derecha) y otro que contiene el complejo (GroEL/GroES):LBL (pico de la izquierda en la **Figura 4.19B**). La comparación del patrón electroforético correspondiente a la fracción soluble (*carril 4* de la **Figura 4.18**) con su perfil de elución cromatográfico en Superdex 200 PG (**Figura 4.19B**) no permite concluir la presencia de LBL no asociada a (GroEL/GroES), aunque en el caso de existir ésta se encontraría formando oligómeros de alto peso molecular ( $> \sim 600$  kDa; rango de separación: 10-600 kDa) o en forma de agregados inespecíficos solubles. La presencia de una banda de aproximadamente 45 kDa en la fracción insoluble (*carril 1* de la **Figura 4.18**) sugiere, de hecho, que LBL recombinante precipita y por ello la banda equivalente observada en la fracción soluble se correspondería con LBL asociada a (GroEL/GroES).



**Figura 4.19.** Estudio de la formación del complejo irreversible con la chaperona GroEL/GroES. **A.** Perfil de elución de LBL expresada sin sobreexpresión conjunta de la chaperona de *E. coli*. **B.** Perfil de elución de LBL coexpresada junto con la chaperona de *E. coli*. En ambos casos se utilizó una columna de exclusión por tamaño Superdex 200 PG preequilibrada en tampón TSA.



**Figura 4.20.** Análisis mediante espectrometría de masas de la formación del complejo GroEL:LBL. En **A** y **B** se muestran los picos obtenidos a partir del análisis del pico que eluyó junto con el volumen de exclusión en la cromatografía de exclusión por tamaño. En **C**, sin embargo, se muestra el espectro obtenido a partir del pico que mostró mayor tiempo de retención en la misma cromatografía.

El análisis de los picos obtenidos en la cromatografía de exclusión por tamaño mediante espectrometría de masas MALDI-TOF confirmó la identidad de las especies moleculares (**Figura 4.20**). El pico del cromatograma de mayor masa molecular (**Figura 4.19B**) se corresponde con dos especies de acuerdo con el espectro de masas: un pico de 57.194,1 Da, correspondiente a GroEL, y otro de 46.689,8 Da, equivalente a la masa de la lectina LBL completa sin la metionina inicial. El segundo pico del cromatograma (el de menor masa) dio lugar también a dos especies con masas asignadas de 19.700,9 y 19.541,6 Da. En principio, ambos picos podrían corresponderse con las especies truncadas LBL<sub>179</sub> y LBL<sub>177</sub>, respectivamente, ambas sin la metionina inicial.

Aunque hay trabajos en los que se describe la obtención de proteínas recombinantes solubles a partir de complejos con GroEL/GroES como los aquí identificados (Goloubinoff *et al.*, 1989; Kolaj *et al.*, 2009), estos pueden considerarse casos excepcionales (Structural Genomics *et al.*, 2008). Además, la aparición de GroEL/GroES después de un proceso de purificación mediante cromatografía de afinidad no es tampoco un resultado extraño, ya que con frecuencia se observa cómo la chaperona suele unirse a matrices de GST, por ejemplo, a través de interacciones inespecíficas o mediante la unión a la proteína de interés (Thain *et al.*, 1996). Existen diversas estrategias que contemplan la adición de GroES y/o ATP para conseguir que la proteína se pliegue adecuadamente (Thain *et al.*, 1996). Sin embargo, los resultados no suelen ser satisfactorios (Katayama *et al.*, 2009). En nuestro caso, la formación del complejo era completamente irreversible, pues ninguna de las estrategias llevadas a cabo consiguió su disociación.

### 2.3.3. Producción de la lectina de *L. bicolor* en *Saccharomyces cerevisiae*

Una evidente limitación de *E. coli* como sistema de expresión recombinante es la ausencia de modificaciones postraduccionales. En algunas ocasiones, la mayor complejidad de las proteínas eucariotas reside en modificaciones covalentes que tienen lugar durante o después del proceso de traducción. Dentro de estas modificaciones es frecuente observar reacciones de acilación, fosforilación, isoprenilación, glicosilación, etc. La N-glicosilación, concretamente, es uno de los tipos de modificaciones postraduccionales más comunes en la naturaleza y mejor estudiados. Sin embargo, la bacteria *E. coli* carece de la maquinaria necesaria para llevar a cabo estas modificaciones (Sethuraman y Stadheim, 2006).

Considerando que algunas lectinas fúngicas muestran un porcentaje elevado de glicosilación, como por ejemplo la lectina de *Rigidoporus lignosus*, cuyo contenido en azúcares representa un 30% de su masa total (Guillot y Konska, 1997) y que la falta de este tipo de

modificación en glicoproteínas suele conducir a su plegamiento incorrecto (Li *et al.*, 2011), analizamos la posibilidad de que LBL tuviese sitios de glicosilación y en ese caso expresarla en sistemas eucariotas. La N-glicosilación consiste en la transferencia de un oligosacárido a la cadena lateral de una asparagina, localizada en la secuencia consenso N-X-S/T, donde la X es cualquier aminoácido excepto Pro.

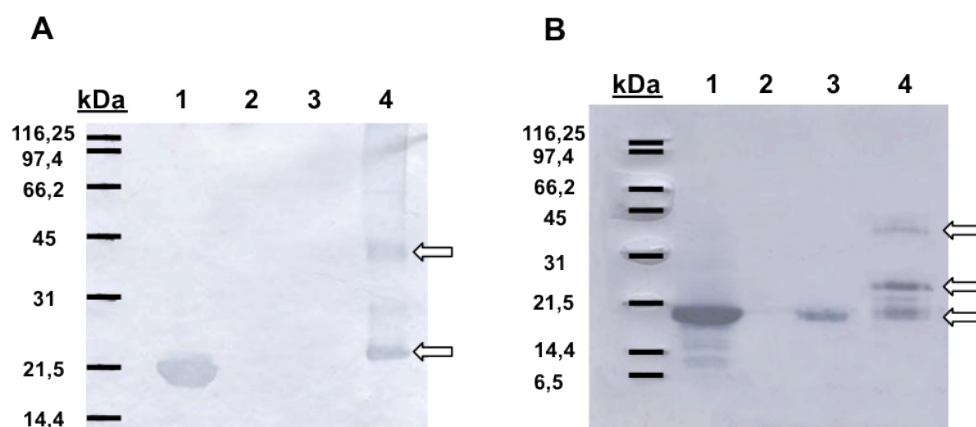
El análisis bioinformático de la secuencia de LBL con el servidor NetNGlyc 1.0 Server ([www.cbs.dtu.dk/services/NetNGlyc](http://www.cbs.dtu.dk/services/NetNGlyc)) predijo seis sitios potenciales de N-glicosilación, los cuales se encuentran localizados en la región C-terminal de la proteína. Aunque no se identificó ningún péptido señal en la secuencia de LBL, actualmente no se sabe ni cómo se reorganiza la secuencia corriente arriba (*upstream*) del gen *lblxp* en el genoma ni cuál podría ser la secuencia consenso de este péptido señal en el organismo *L. bicolor*.

Teniendo en cuenta estos resultados, se expresó la lectina LBL en *Saccharomyces cerevisiae*, microorganismo que tiene la capacidad de realizar modificaciones postraduccionales, como la N-glicosilación (Kingsman *et al.*, 1985). Es cierto, no obstante, que *S. cerevisiae* no introduce los mismos azúcares en las modificaciones que otros organismos superiores, sino que muestra predilección por la incorporación de residuos de manosa (Aebi, 2013).

De este modo, el gen *lblxp* fue clonado en el vector YE<sub>p</sub>FLAG, que contiene la secuencia del péptido señal FLAG que permite exportar la proteína al medio extracelular, pasando previamente por el retículo endoplásmico, donde se producen las modificaciones postraduccionales pertinentes. Después de la inducción de las levaduras, éstas se separaron del medio extracelular, que fue utilizado posteriormente para la purificación de la lectina. La presencia de LBL en el medio extracelular se analizó mediante inmunodetección con anticuerpos anti-LBL<sub>152</sub> (sección 3.7 de *Materiales y Métodos*).

El *western blot* (**Figura 4.21A**) muestra que la fracción purificada mediante cromatografía de afinidad a Sepharose® 4B está constituida principalmente por una mezcla de dos especies. La banda mayoritaria observada se correspondería de nuevo con una forma truncada de la lectina, cuya masa molecular está cerca de los 21 kDa. Este resultado coincide con lo observado previamente en *E. coli*. Por otro lado, aparece una banda más tenue cerca de los 45 kDa, que podría pertenecer a la lectina completa sin azúcares, indicando la ausencia de glicosilación. En cualquier caso, la existencia de hidrólisis de LBL, y el bajo rendimiento en su producción, indica que esta estrategia de expresión-purificación no ofrece ventajas sustanciales con el uso de *E. coli* para llevar a cabo estudios cristalográficos.

Teniendo en cuenta que el resultado anterior apunta a que no existe glicosilación sobre LBL y con la intención de comprobar si el medio intracelular de *S. cerevisiae* es más beneficioso para la producción y el plegamiento de LBL que el ofrecido por *E. coli*, la lectina se expresó sin el péptido FLAG. En este caso, las levaduras se lisaron y fue la fracción soluble del extracto celular la que se purificó mediante cromatografía de afinidad a Sepharose® 4B. El análisis de la fracción purificada vuelve a darnos el mismo resultado que los obtenidos anteriormente (**Figura 4.21B**). La proteína LBL experimenta una proteólisis que le lleva a producir principalmente el módulo lectina N-terminal de forma aislada. De hecho, se observó que cuando la muestra se deja en nevera un tiempo prolongado, la fracción se homogeniza y en el *western blot* sólo se visualiza la banda correspondiente al módulo lectina.

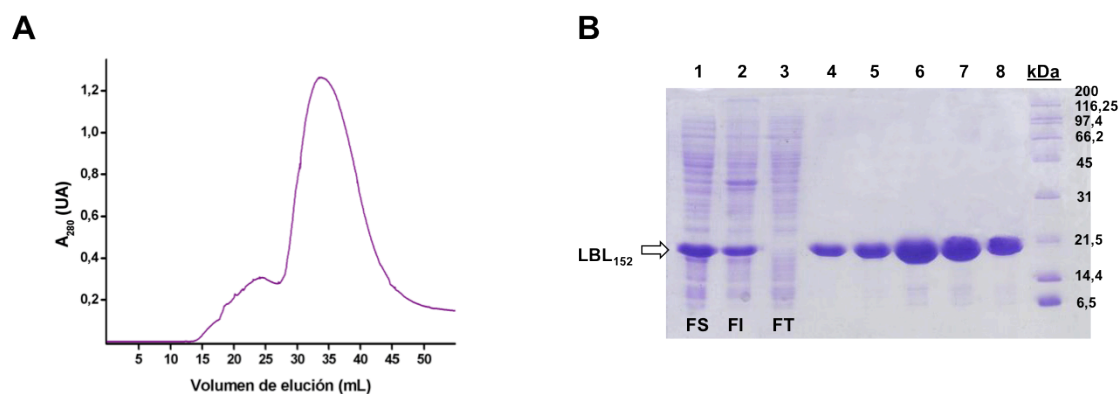


**Figura 4.21.** Análisis mediante *western blot* de la producción de LBL usando *S. cerevisiae* como sistema de expresión heterólogo. **A.** Producción extracelular de LBL. Carril 1: control positivo con el módulo lectina N-terminal de LBL (LBL<sub>152</sub>). Carril 2: medio extracelular. Carril 3: fracción no retenida. Carril 4: fracción purificada mediante cromatografía de afinidad en Sepharose® 4B. **B.** Producción intracelular de LBL. Carril 1: control positivo con el módulo lectina N-terminal de LBL (LBL<sub>152</sub>). Carril 2: fracción insoluble. Carril 3: fracción purificada tras una semana en la nevera (4 °C). Carril 4: fracción recién purificada. La inmunodetección se llevó a cabo usando como anticuerpo primario un anticuerpo anti-LBL<sub>152</sub> de conejo. Las flechas indican la migración de las especies que reaccionan positivamente con los anticuerpos.

Los resultados obtenidos indican que ni *E. coli* ni *S. cerevisiae* son sistemas eficaces de expresión para obtener la proteína LBL de forma recombinante. Una dificultad añadida en el estudio de la proteína completa reside en el hecho de que su región C-terminal no presenta similitud de secuencia con ninguna otra proteína caracterizada hasta el momento, tal y como se observa al llevar a cabo una búsqueda de homólogos (BLAST, ExPASy). Teniendo en cuenta que el objetivo principal de este trabajo es caracterizar los módulos lectina con estructura tipo trébol  $\beta$ , decidimos continuar el estudio centrándonos únicamente en el dominio N-terminal de la lectina de *L. bicolor*.

## 2.4. Producción y purificación del módulo lectina N-terminal de LBL

La transformación del plásmido pKLBL<sub>152</sub> en células de *E. coli* BL21(DE3) permitió expresar el correspondiente módulo lectina N-terminal de la proteína LBL (LBL<sub>152</sub>). Éste se purificó a homogeneidad mediante una cromatografía de afinidad con Sepharose® 4B y posterior elución con 0,2 M lactosa. Las fracciones recogidas fueron analizadas mediante SDS-PAGE (Figura 4.22). Teniendo en cuenta que la masa molecular de LBL<sub>152</sub> es de 16,8 kDa, los resultados indican que el módulo lectina se purificó en un único paso, mostrándose una elevada tasa de producción: unos 110 mg/L de cultivo. Este resultado es consistente con lo observado anteriormente en LSL<sub>150</sub> y en otras lectinas fúngicas, ya que el valor se encuentra dentro del mismo orden de magnitud, tal y como se ha discutido más arriba en el apartado 1.1 de *Resultados y Discusión*.



**Figura 4.22.** Purificación del módulo lectina N-terminal LBL<sub>152</sub>. **A.** Perfil de elución de la purificación de LBL<sub>152</sub> mediante cromatografía de afinidad a Sepharose® 4B. **B.** SDS-PAGE (15% de poliacrilamida) de la purificación de LBL<sub>152</sub> a partir de la elución de la columna de Sepharose® 4B. Carril 1: fracción soluble del extracto celular (FS). Carril 2: fracción insoluble del extracto celular (FI). Carril 3: fracción no retenida o *flow-through* (FT). Los carriles 4-8 se corresponden con el máximo observado en el perfil de elución.

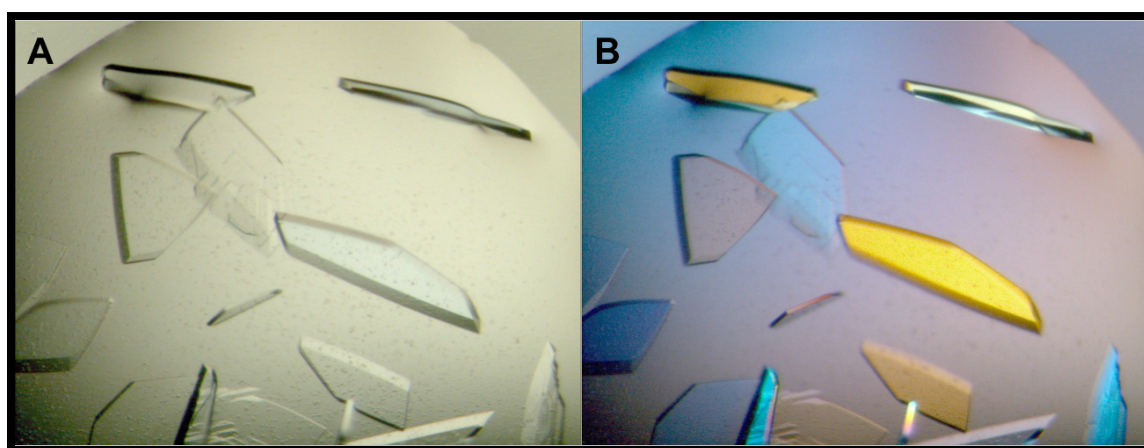
Al igual que LSL<sub>150</sub>, LBL<sub>152</sub> se produce eficazmente en *E. coli*, lo que indica que el correspondiente fragmento génico tiene una elevada capacidad para traducirse y que el entorno ofrecido por la bacteria es ideal para el correcto plegamiento de la proteína. Este comportamiento también se observa en el caso de la producción de la lectina completa, puesto que la banda correspondiente a la región N-terminal siempre presenta una alta intensidad, indicando su elevada expresión. Con respecto a su función como un agente solubilizante natural, la producción de LBL no nos ha aportado pistas acerca de ello debido a los problemas de plegamiento intrínsecos que se detectaron. Sin embargo, sabiendo que LSL<sub>150</sub> es un modelo que muestra bastante similitud, se podría pensar que el módulo N-terminal de la lectina de *L. bicolor* ejerciese una función parecida.



## 2.5. La estructura de LBL<sub>152</sub> revela a un nuevo trébol $\beta$

### 2.5.1. Cristalización de LBL<sub>152</sub>

Teniendo en cuenta la concentración a la que cristalizó LSL<sub>150</sub> (apartado 1.2.1 de *Resultados y Discusión*), LBL<sub>152</sub> se concentró inicialmente hasta 22 mg/mL. Normalmente, imitar las condiciones de cristalización de proteínas homólogas puede dar buen resultado (Garman y Garboczi, 2004). A esta concentración se realizaron los primeros ensayos de cristalización a 18 °C, usando el método de difusión de vapor en gota sentada, testándose así cerca de 600 condiciones distintas. Los primeros resultados de este *screening* fueron todos negativos. La mayor parte de las gotas permanecían claras y sólo se apreciaba una baja proporción de precipitados amorfos. Los mismos ensayos de cristalización se repitieron incrementando la concentración de proteína. A pesar de aumentar la proporción de precipitados en las gotas, a 90 mg/mL prácticamente el 50% de la gotas seguían permaneciendo claras.



**Figura 4.23.** Cristales de LBL<sub>152</sub> en presencia de 0,2 M lactosa obtenidos en 20% (p/v) PEG 6000, 100 mM acetato de sodio pH 5,0 y 200 mM cloruro de amonio. Se muestra la misma gota bajo luz blanca (A) y bajo luz polarizada (B).

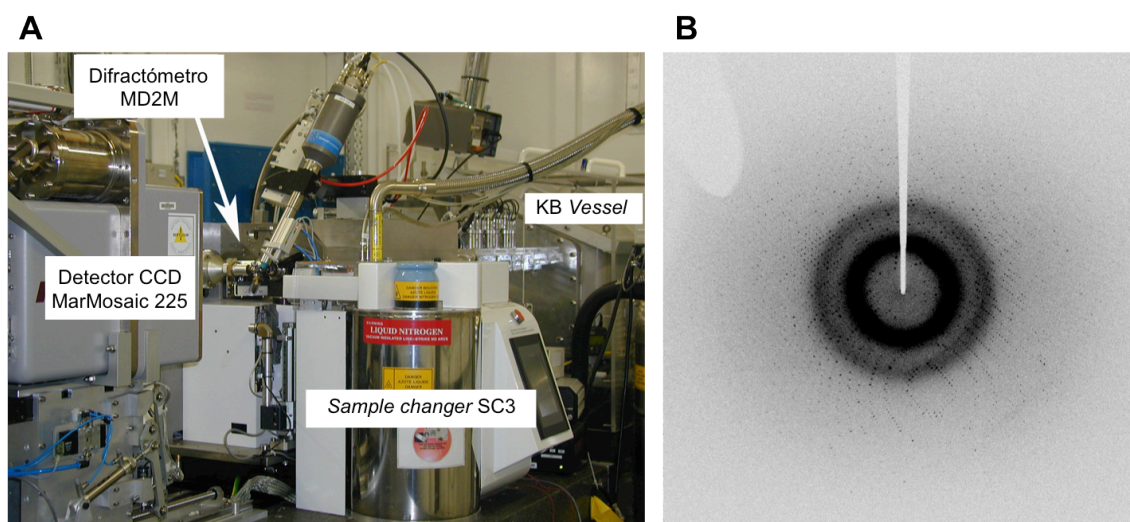
A la vista del comportamiento marcadamente diferente entre LSL<sub>150</sub> y LBL<sub>152</sub> a nivel de cristalización, se optó por llevar a cabo experimentos en presencia de ligandos, en particular, lactosa con la que debería interaccionar de acuerdo a su comportamiento cromatográfico sobre Sepharose® 4B. En este sentido, la adición de ligandos puede hacer más compacta a la proteína y aumentar, por tanto, la probabilidad de que cristalice (Russo Krauss *et al.*, 2013). De este modo, llevamos a cabo un ensayo de cristalización a pequeña escala en el que probamos las mismas 600 condiciones iniciales, pero utilizando una muestra de LBL<sub>152</sub> a 24 mg/mL preincubada durante ~2 horas con 0,2 M lactosa. Al cabo de 3-4 horas se obtu-

vieron cristales en una gran cantidad de condiciones de cristalización diferentes con PEG como agente precipitante y pH ácido.

Después de un proceso de optimización, los mejores cristales (**Figura 4.23**) se obtuvieron con el sistema de difusión de vapor en gota sentada, mezclando 1  $\mu\text{L}$  de LBL<sub>152</sub> en tampón TSAL con 1  $\mu\text{L}$  de solución precipitante: 20% (*p/v*) PEG 6000, 0,1 M acetato de sodio pH 5,0 y 0,2 M cloruro de amonio. Estos adoptan una forma de tipo púa, alcanzando su tamaño final al cabo de una semana. Los cristales se caracterizan por ser birrefringentes, indicando una estructura interna bien ordenada (Luft *et al.*, 2011).

### 2.5.2. Determinación estructural de LBL<sub>152</sub> en complejo con lactosa

Antes de medir los cristales optimizados de LBL<sub>152</sub>, estos se crioprotegieron con la misma solución utilizada para cristalizar suplementada con glicerol al 20% (*v/v*). Los cristales fueron medidos posteriormente en la línea ID23-2 del ESRF (Grenoble, Francia) a una temperatura constante de 100 K. Esta línea trabaja a una longitud de onda fija de 0,873 Å (14,20 keV) y posee un detector de tipo CCD MarMosaic 225. Se recogieron un total de 200 imágenes con un ángulo de giro por imagen de 1°. Las imágenes recogidas revelan un patrón de difracción que indica que los cristales tienen muy buena calidad, como se deduce al observar reflexiones muy nítidas que alcanzan una resolución de  $\sim 1$  Å (**Figura 4.24**).



**Figura 4.24.** A. Interior del *hutch* de experimentación de la línea ID23-2 del ESRF (Grenoble, Francia). B. Patrón de difracción de los cristales de LBL<sub>152</sub> en presencia de 0,2 M lactosa, usando 20% (*v/v*) glicerol como agente crioprotector.

El análisis de los datos cristalográficos, llevado a cabo con iMOSFLM, indicó que el cristal es ortorrómbico, perteneciendo al grupo espacial  $P2_12_12$  cuya celda unidad tiene los siguien-

tes parámetros:  $a = 52,53 \text{ \AA}$ ,  $b = 61,27 \text{ \AA}$  y  $c = 44,79 \text{ \AA}$ . La unidad asimétrica está constituida por una única molécula de proteína y el cristal posee un porcentaje de solvente del 42,6%, determinado a partir del volumen de Matthews (Kantardjieff y Rupp, 2003). Después de las etapas de integración y escalado, el análisis estadístico reflejó que el set de datos posee una calidad excepcional, presentando una completitud del 100% hasta una resolución de  $1 \text{ \AA}$  (resolución atómica).

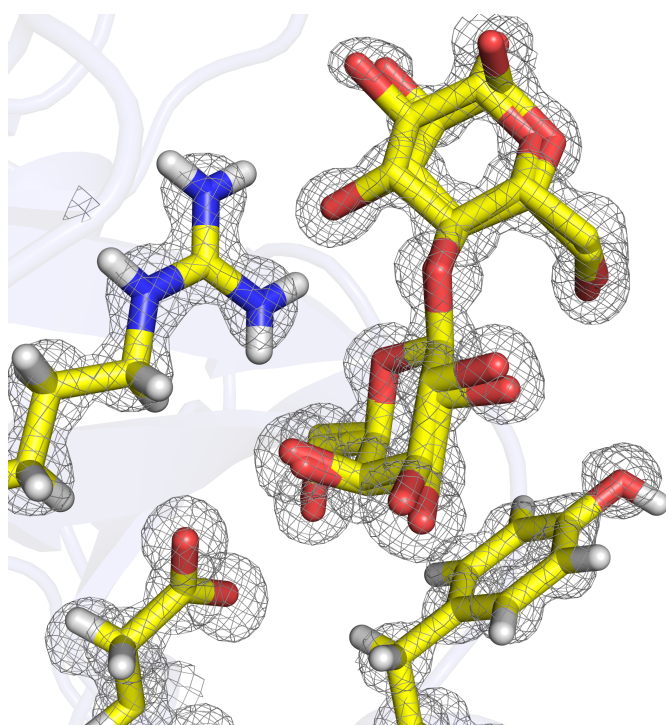
Parámetros	LBL <sub>152</sub> :(lactosa) $\alpha\beta\gamma$
Recogida de datos y procesamiento	
Grupo espacial	$P2_12_12$
Moléculas en la UA	1
Celda unidad ( $\text{\AA}$ )	52,53, 61,27, 44,79
Longitud de onda ( $\text{\AA}$ )	0,873
Resolución ( $\text{\AA}$ )	15,2–0,99
Reflexiones totales	373.565
Reflexiones únicas	79.595
Completitud	100 (99,8)
Multiplicidad	8,0 (7,7)
$R_{\text{merge}}$ (%)	11,5 (35,3)
$I/\sigma(I)$	10,9 (5,1)
Refinamiento	
$R_{\text{work}}/R_{\text{free}}$	0,126/0,139
RMSD de los enlaces ( $\text{\AA}$ )	0,009
RMSD de los ángulos ( $^\circ$ )	1,617
Factor $B$ medio ( $\text{\AA}^2$ )	9,09
Número de átomos	
Proteína	1.670
Hidrógenos	1.198
Lactosa	69
Agua	279
Mapa de Ramachandran	
Favorecidos (%)	96,2
Permitidos (%)	3,2
Prohibidos (%)	0,6

\*Los valores de la concha de mayor resolución se muestran en paréntesis.

**Tabla 4.4.** Datos de difracción y estadísticas del refinamiento de la estructura de LBL<sub>152</sub> en complejo con lactosa (LBL<sub>152</sub>:lactosa).

La estructura de LBL<sub>152</sub> se resolvió con PHASER (McCoy *et al.*, 2007) mediante reemplazo molecular, utilizando como modelo de reemplazo la estructura del módulo lectina LSL<sub>150</sub>. El modelo final se obtuvo tras varios ciclos de refinamiento utilizando la herramienta phenix.refine del programa PHENIX (Adams *et al.*, 2010). Cuando la resolución es muy elevada y, por lo tanto, el número de observaciones experimentales es grande, se recomienda aplicar modelos de refinamiento no isótropos que consideren desplazamientos o vibraciones de tipo elipsoide (*thermal ellipsoid*) con respecto de la posición de equilibrio de los átomos (Vrielink y Sampson, 2003; Merritt, 2012). En nuestro caso, aplicamos un refi-

namiento modelando dichos desplazamientos anisótropos, que suponen la estimación de seis parámetros o ADPs (*anisotropic displacement parameters*) por átomo, además de los tres correspondientes a su posición de equilibrio (Merritt, 2012). Igualmente, se procedió a la incorporación de átomos de hidrógeno a la estructura. En cristalografía de rayos X, los hidrógenos no suelen incluirse en el modelo final ya que sus posiciones no pueden determinarse debido a que son muy débiles dispersores de rayos X. Sin embargo, a resoluciones muy altas, el mapa Fourier diferencia ( $F_o - F_c$ ) puede revelar la posición de estos (Vrielink y Sampson, 2003; Phenix, septiembre de 2013). Los átomos de hidrógeno sólo se incluyeron en la molécula de proteína, no en los ligandos. Por último, durante los ciclos de reconstrucción de la cadena polipeptídica se pudieron identificar 15 residuos que presentaban dos conformeros definidos. Los valores finales de  $R_{work}$  y  $R_{free}$  fueron 12,6 y 13,8%, respectivamente.



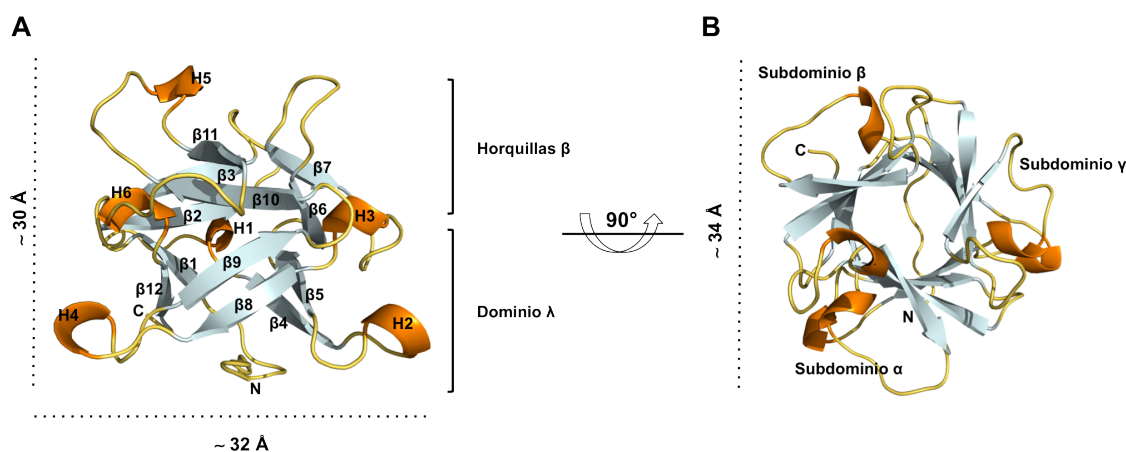
**Figura 4.25.** Vista detallada de uno de los sitios de unión de LBL<sub>152</sub>, donde se aprecia la calidad del mapa de densidad electrónica  $2F_o - F_c$  (gris) a  $1\sigma$ . La elevada resolución permitió incorporar átomos de hidrógeno (blanco) en la estructura de la lectina.

El modelo atómico de LBL<sub>152</sub> está compuesto por 1670 átomos (más 1198 hidrógenos), perfectamente definidos en densidad (**Figura 4.25**), 69 átomos correspondientes a 3 moléculas de lactosa y 279 moléculas de solvente. Las estadísticas de toma de datos y refinamiento se pueden ver en la **Tabla 4.4**. El modelo revela que la lectina une un total de tres moléculas de ligando, una por cada uno de los subdominios ( $\alpha$ ,  $\beta$  y  $\gamma$ ). Este resultado marca una de las principales diferencias con su homólogo LSL<sub>150</sub>, como se verá más adelante en el apartado 2.5.4.



### 2.5.3. Descripción de la estructura de LBL<sub>152</sub>

El módulo lectina LBL<sub>152</sub> presenta un plegamiento trébol  $\beta$ , similar al descrito anteriormente para LSL<sub>150</sub> (apartado 1.2.2 de *Resultados y Discusión*), con pseudosimetría interna de orden 3, resultado de la aparición de tres subdominios estructuralmente similares ( $\alpha$ ,  $\beta$  y  $\gamma$ ) (**Figura 4.26**). Cada subdominio, con  $\sim 40$ -50 residuos, está formado por cuatro hebras  $\beta$ . La primera y la cuarta hebra de cada subdominio ( $\beta 1$ ,  $\beta 4$ ,  $\beta 5$ ,  $\beta 8$ ,  $\beta 9$  y  $\beta 12$ ) están implicadas en la formación del barril  $\beta$  (dominio  $\lambda$ ), mientras que la segunda y la tercera ( $\beta 2$ ,  $\beta 3$ ,  $\beta 6$ ,  $\beta 7$ ,  $\beta 10$  y  $\beta 11$ ), forman parte del triplete de horquillas  $\beta$ , donde se organizan los sitios de unión de azúcares. Esta distribución hace que los extremos N y C-terminales estén en uno de los extremos del barril  $\beta$ , mientras que los sitios de unión de azúcares se encuentran en el extremo opuesto. En el caso particular de LBL<sub>152</sub>, cada subdominio está intercalado con una corta hélice  $3_{10}$  (H2 y H4), que no aparece en otros tréboles  $\beta$  ya resueltos, excepto en el módulo lectina N-terminal de LSLa (Mancheño *et al.*, 2005). Además, en los tres subdominios, entre la tercera y la cuarta hélice también se disponen unas pequeñas hélices del mismo tipo (H1, H3 y H6). En el subdominio  $\gamma$ , en cambio, se localiza otra hélice  $3_{10}$  extra entre la segunda y la tercera hebra  $\beta$  (H5).

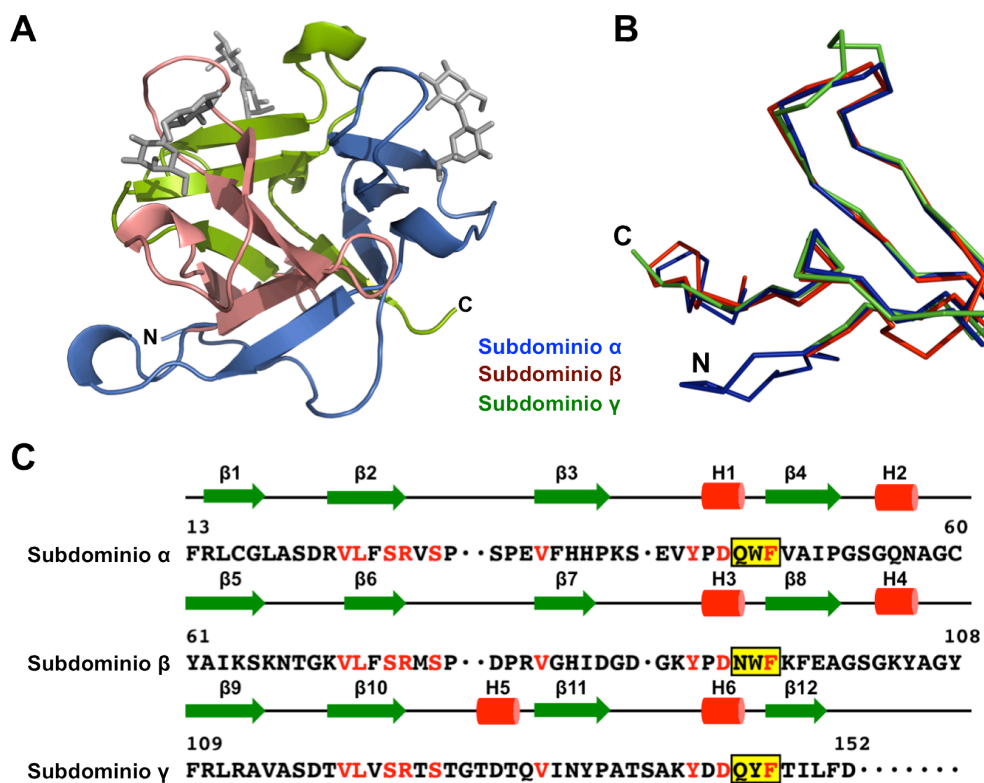


**Figura 4.26.** Estructura cristalográfica del módulo lectina N-terminal LBL<sub>152</sub>. En ambas imágenes se detallan los elementos de estructura secundaria: hebras  $\beta$  (cian) y hélices  $3_{10}$  (naranja). **A.** Visión frontal de LBL<sub>152</sub> donde se representa el dominio  $\lambda$  y el triplete de horquillas  $\beta$ . Las hebras  $\beta$  y las hélices  $3_{10}$  aparecen numeradas sobre la estructura. **B.** Visión cenital de LBL<sub>152</sub> donde se identifican los tres subdominios que componen la estructura. Las dimensiones aproximadas del trébol  $\beta$  se muestran en los laterales, dándose las longitudes en ångströms (Å).

LBL<sub>152</sub> tiene unas dimensiones aproximadas de  $30 \times 32 \times 34 \text{ Å}^3$  y un área expuesta al solvente de  $\sim 7300 \text{ Å}^2$ . El análisis del empaquetamiento cristalino mediante el servidor PISA (Krissinel y Henrick, 2007) descarta la posibilidad de formación de oligómeros, ya que las

áreas de contacto son pequeñas ( $\sim 400\text{-}680 \text{ \AA}^2$ ), representando menos del 10% del área expuesta al solvente. Se observa la existencia de contactos cristalinos mediados por las moléculas de lactosa, que siempre aparecen unidas en los sitios de unión del trébol  $\beta$ , y que podrían explicar la necesidad de este azúcar para cristalizar LBL<sub>152</sub>.

Si se compara la estructura de LBL<sub>152</sub> con la de los otros tréboles  $\beta$  fúngicos ya caracterizados estructuralmente (ver apartado 5.2 de la *Introducción*), se observa que la estructura global se conserva: el valor medio de los RMSD de los C $\alpha$  de los residuos en posiciones equivalentes es de tan sólo 1,94 Å. Esto contrasta con las estructuras primarias en donde a excepción de la elevada similitud que presenta con la secuencia de LSLa (52%), los porcentajes determinados con el resto de lectinas fúngicas es muy bajo, no superando el 25% en ningún caso (a excepción de la proteína BEL de *B. edulis* con un 43% de identidad). Este resultado era esperable en el contexto evolutivo de los tréboles  $\beta$  que considera su aparición como un acontecimiento recurrente al cual le sigue una importante divergencia de secuencia (Broom *et al.*, 2012).



**Figura 4.27.** Pseudosimetría interna en LBL<sub>152</sub>. **A.** Diagrama de cintas de la estructura de LBL<sub>152</sub> en la que se muestran los subdominios:  $\alpha$  (azul),  $\beta$  (rojo) y  $\gamma$  (verde). Las moléculas de lactosa aparecen sobre cada uno de los sitios de unión de azúcares (representadas con modelo de varillas y en color gris). **B.** Superposición de la cadena de átomos C $\alpha$  de cada uno de los subdominios de LBL<sub>152</sub>. **C.** Disposición de los elementos de estructura secundaria a lo largo de la secuencia. Las hebras  $\beta$  están representadas como flechas verdes y las hélices  $\alpha$ , como cilindros rojos. La secuencia consenso de los tréboles  $\beta$  (QXW)<sub>3</sub> está indicada con un recuadro amarillo. Además, en rojo se representan los residuos conservados entre los tres subdominios de la secuencia.

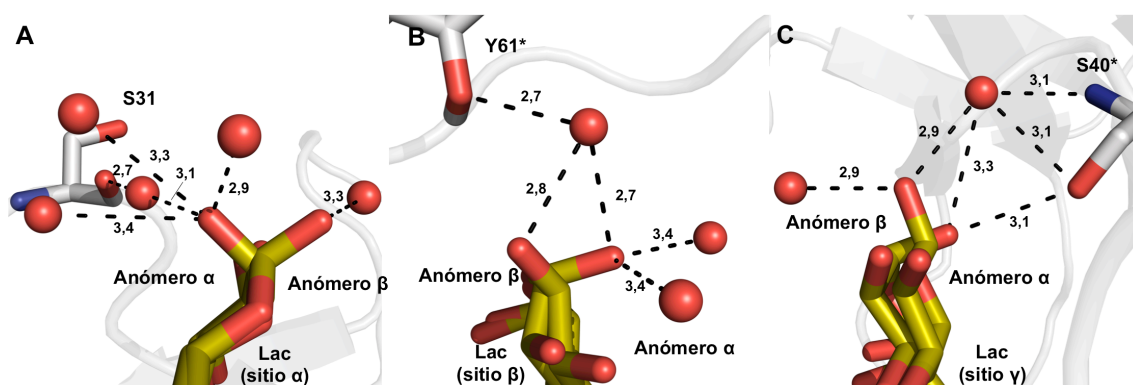
Al comparar la estructura y la secuencia de los tres subdominios, ocurre algo similar a lo descrito arriba para los tréboles  $\beta$  entre sí (**Figura 4.27**). Desde un punto de vista estructural, los tres subdominios son prácticamente superponibles, con valores de RMSD (de todos los átomos) de: 0,98 Å entre  $\alpha$  y  $\beta$ ; 0,54 Å entre  $\alpha$  y  $\gamma$ ; y 0,51 Å entre  $\beta$  y  $\gamma$ . La similitud entre las secuencias es más baja, siendo del 41,7% entre  $\alpha$  y  $\beta$ ; 41,5% entre  $\alpha$  y  $\gamma$ ; y 26,8% entre  $\beta$  y  $\gamma$ . Con respecto a la secuencia, el típico triplete (QXW)<sub>3</sub> (Hazes, 1996) aparece también en LBL<sub>152</sub>. Sin embargo, éste presenta algunas singularidades ya que en el sitio  $\beta$  aparece una Asn en vez de una Gln y en los tres subdominios, el Trp ha sido sustituido por otro residuo aromático, una Phe.

#### 2.5.4. Estudio detallado del modo de unión de lactosa

La alta resolución de los datos cristalográficos y la buena calidad del mapa de densidad electrónica han permitido analizar en detalle las interacciones que se establecen entre los residuos de los sitios de unión del módulo lectina LBL<sub>152</sub> y las moléculas de ligando y solvente. En primer lugar, al igual que ocurría con las moléculas de lactosa unidas a LSL<sub>150</sub>, el modelo muestra cómo ambos anómeros de la lactosa se unen a los tres sitios de unión disponibles. El mapa inicial ( $F_o - F_c$ ) en torno al oxígeno O1' del anillo de glucosa indicaba que ambas orientaciones, axial y ecuatorial, podían ser modeladas. Las ocupaciones se ajustaron manualmente en función de la señal procedente del mapa diferencia. En los tres casos, los valores del ratio anómero  $\alpha/\beta$  referente a las ocupaciones obtenidas son muy similares entre sí, siendo de 0,39/0,61, 0,55/0,45 y 0,27/0,73 para los sitios  $\alpha$ ,  $\beta$  y  $\gamma$ , respectivamente. La comparación de estos resultados con la distribución de ambos ánomeros para la glucosa en solución (0,38/0,62) (Vyas *et al.*, 1994) permiten deducir que no existe una clara preferencia por un anómero u otro, por lo que LBL<sub>152</sub> tampoco presentaría una especificidad anomérica.

El análisis de la estructura revela interacciones que implican al oxígeno O1' del carbono anomérico (**Figura 4.28**). En relación con la lactosa del sitio  $\alpha$  (**Figura 4.28A**), mientras que el anómero  $\beta$  sólo establece un enlace de hidrógeno con una molécula de agua, el anómero  $\alpha$  tiene capacidad para interactuar con cuatro moléculas de agua. Además, una de ellas interactúa a su vez mediante un enlace de hidrógeno con el oxígeno carbonílico de la cadena principal del residuo de Ser-31 de LBL<sub>152</sub>. En el sitio  $\beta$  la situación es inversa (**Figura 4.28B**): en este caso, el oxígeno O1' del anómero  $\alpha$  interactúa con tres moléculas de agua y el del anómero  $\beta$  solamente con una. La molécula de agua que está unida a ambos

anómeros forma un enlace de hidrógeno con la Tyr-61 que, en este caso, pertenece a una de las moléculas relacionadas por simetría. Por último, con respecto al sitio  $\gamma$ , el anómero  $\beta$  forma dos enlaces de hidrógeno con dos moléculas de agua y el anómero  $\alpha$ , por otro lado, enlaza con una de las moléculas de agua anteriores y con el residuo de Ser-40, que también proviene de una molécula de LBL<sub>152</sub> relacionada por simetría.



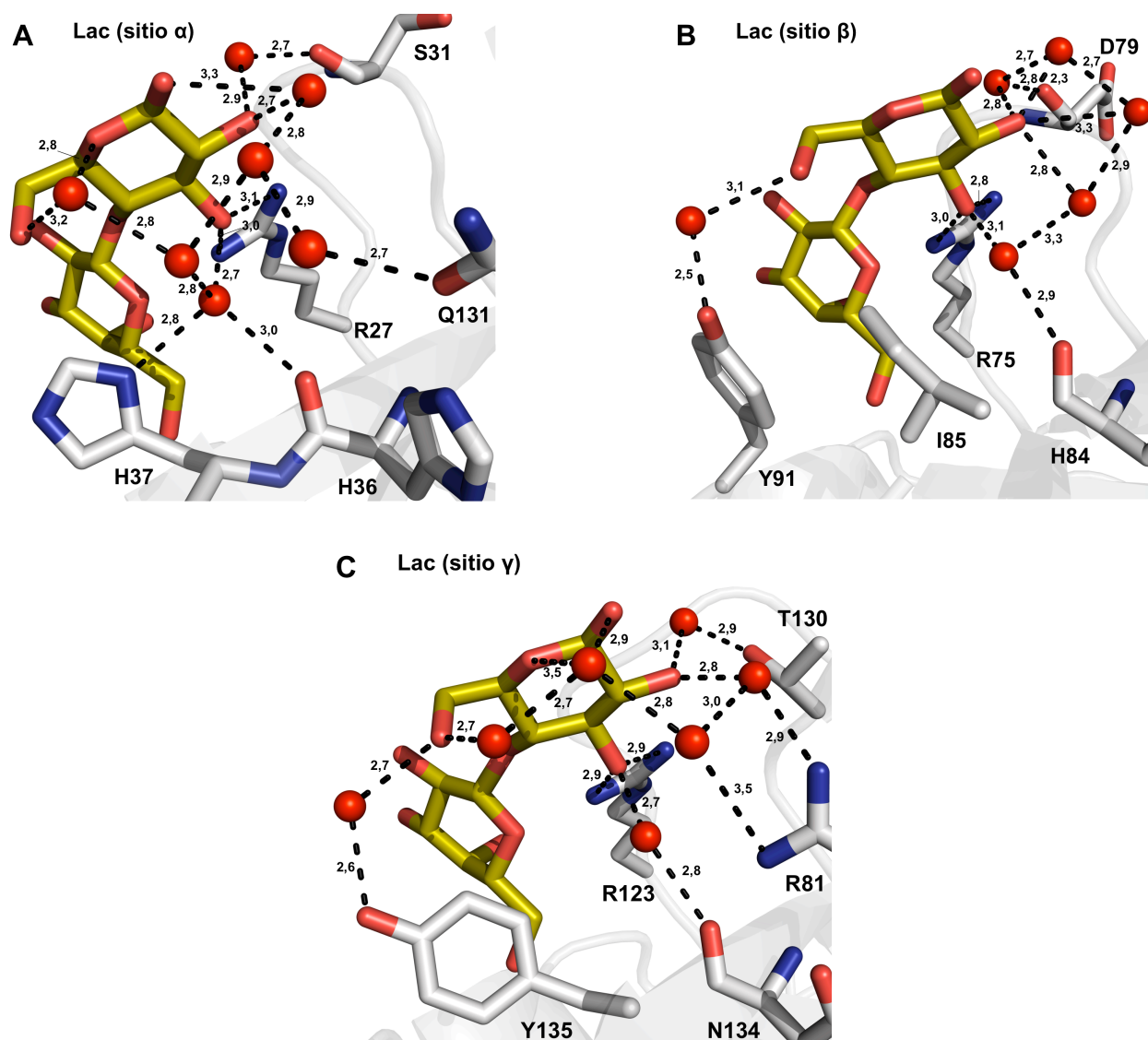
**Figura 4.28.** Vista detallada de las interacciones específicas entre los oxígenos del carbono anomérico de ambos anómeros  $\alpha$  y  $\beta$  de la lactosa y residuos de la molécula LSL<sub>150</sub>. En **A** se muestran las interacciones del sitio  $\alpha$ ; en **B**, las del sitio  $\beta$ ; y en **C**, las del sitio  $\gamma$ . Las moléculas de agua se representan mediante *esferas* de color *rojo*. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno que se encuentran dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea de puntos* y las medidas de las distancias vienen dadas en *ångströms* (Å). El asterisco (\*) indica que los aminoácidos pertenecen a moléculas de proteína relacionadas por simetría.

A la vista de estos resultados, se puede hacer una consideración interesante: el hecho de que las ocupaciones de los anómeros  $\alpha/\beta$  de lactosa unida a cualquiera de los sitios de unión de LBL<sub>152</sub> coincida razonablemente bien con la relación de los mismos en solución, aun cuando en el empaquetamiento cristalino el átomo de oxígeno O1' establece distintas interacciones en un caso u otro, implica que estas interacciones esencialmente no condicionan el empaquetamiento o lo hacen de un modo equivalente.

Ambos anómeros de la lactosa comparten las interacciones que se establecen con el resto de átomos de los anillos de glucosa y galactosa porque la única posición cuya conformación varía entre ellos es la del O1'. Con respecto a las interacciones que se establecen con el anillo de glucosa, en los tres sitios de unión aparece una intrincada red de enlaces de hidrógeno mediada por moléculas de solvente (**Figura 4.29**). En primer lugar, comenzando por el sitio  $\alpha$  (**Figura 4.29A**), el átomo de oxígeno O2' forma dos enlaces de hidrógeno con moléculas de agua; una de ellas interacciona con la Ser-31, mientras que la otra participa en una red de enlaces entre diferentes moléculas de agua. Estas moléculas, a su vez, interaccionan con los oxígenos O3', O5' y O6', así como con el nitrógeno N $\delta$ -1 de la His-37, el



oxígeno carbonílico de la cadena principal de la His-36 y el oxígeno O $\epsilon$ -1 de la Gln-131. Adicionalmente, el oxígeno O3' forma dos enlaces de hidrógeno con el residuo de Arg-27.



**Figura 4.29.** Vista detallada de las interacciones específicas entre el anillo de glucosa de la lactosa (se representa solamente el anómero  $\beta$ ) y los residuos del sitio de unión de LBL<sub>152</sub>. En **A** se muestran las interacciones del sitio  $\alpha$ ; en **B**, las del sitio  $\beta$ ; y en **C**, las del sitio  $\gamma$ . Las moléculas de agua se representan mediante *esferas de color rojo*. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno que se encuentran dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea de puntos* y las medidas de las distancias vienen dadas en *ångströms* (Å).

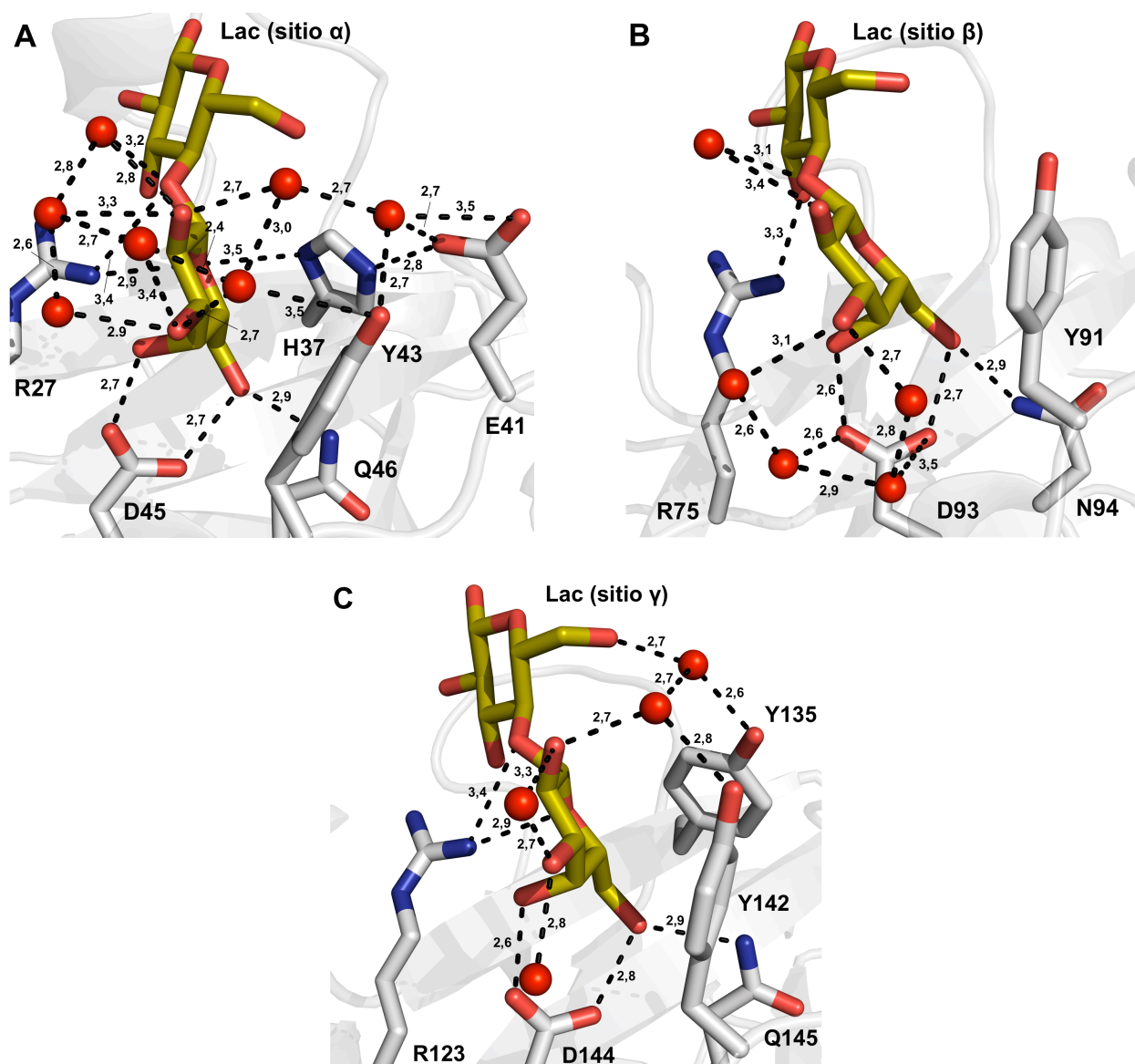
Por otro lado, en el sitio  $\beta$  (**Figura 4.29B**) hay un menor número de interacciones posiblemente por la presencia de la cadena lateral alifática de la Ile-85. Sin embargo, se observa cómo los oxígenos O2' y O3' participan en una red de moléculas de agua que les lleva a interactuar de forma indirecta con los residuos de Asp-79 e His-84. El oxígeno O3', además, al igual que en el anterior caso, interacciona directamente con los nitrógenos N $\eta$

del grupo guanidinio de la Arg-75. El oxígeno O6' también forma un enlace de hidrógeno con una molécula de agua, que reconoce asimismo al grupo hidroxilo de la Tyr-91.

Por último, en el sitio  $\gamma$  (**Figura 4.29C**) vuelve a repetirse la misma situación. El oxígeno del anillo de glucosa junto con los oxígenos O1', O2', O3' y O6' forman una red de enlaces de hidrógeno con moléculas de agua que les acercan a los residuos de Arg-81, Thr-130, Asn-134 y Tyr-135. El oxígeno O3', finalmente, es el único que establece un contacto directo con el aminoácido Arg-123 de forma similar a lo descrito en los dos casos anteriores.

La arquitectura de los tres sitios de unión es muy similar a la que se había descrito previamente para el sitio  $\gamma$  de LSL<sub>150</sub>. Se observa que la función del residuo de Arg que interacciona directamente con el oxígeno O3' del anillo de glucosa es una característica común que se mantiene en los tres sitios de unión de LBL<sub>152</sub> (Arg-27, Arg-75 y Arg-123) y en los dos sitios operativos de LSL<sub>150</sub> (Arg-75 y Arg-123), indicando, por lo tanto, que este aminoácido es importante en el reconocimiento del anillo de glucosa. De hecho, en la lectina BEL, la que mayor identidad de secuencia comparte con LSL<sub>150</sub> y LBL<sub>152</sub>, también se observa la presencia de un residuo de Arg en cada uno de los sitios de unión, que establece contactos con la unidad de glucosa de la molécula de lactosa (Bovi *et al.*, 2013). Este residuo, sin embargo, no se ha identificado en otros tréboles  $\beta$  fúngicos, donde el modo de reconocimiento es diferente. Por ejemplo, el dominio lectina de unión a xilano de la xilanasa de *Streptomyces lividans* (Notenboom *et al.*, 2002) y la lectina I de muérdago (Mikeska *et al.*, 2005) fijan el oxígeno O3' de la glucosa mediante un enlace de hidrógeno formado con un residuo de Asp (Asp-22 y Asp-26, respectivamente), mientras que la lectina EW29 de la lombriz de tierra *Lumbricus terrestris* posee un residuo de Lys en su lugar (Lys-233) (Suzuki *et al.*, 2009).

En último lugar, se van a detallar las interacciones presentes entre el anillo de galactosa y los aminoácidos de los sitios de unión. Como ocurre en la práctica totalidad de las lectinas con plegamiento trébol  $\beta$  y especificidad por  $\beta$ -galactósidos, LBL<sub>152</sub> dispone en sus tres sitios de unión de aminoácidos aromáticos que forman un apilamiento con la cara B del anillo de galactosa (**Figura 4.30**). Esta interacción es fuerte y orienta al disacárido para poder establecer el resto de interacciones intermoleculares. LBL<sub>152</sub> posee residuos de Tyr (Tyr-43, Tyr-91 y Tyr-142 en los sitios  $\alpha$ ,  $\beta$  y  $\gamma$ , respectivamente), lo que es frecuente en los tréboles  $\beta$  de origen fúngico (**Tabla 4.3**). Cabe añadir que los factores *B* de los átomos del anillo de galactosa y de los residuos que conforman el sitio de unión son muy parecidos, lo que indica que todas las moléculas de proteína tienen lactosa unida.



**Figura 4.30.** Vista detallada de las interacciones específicas entre el anillo de galactosa de la lactosa (se representa solamente el anómero  $\beta$ ) y los residuos de unión de LBL<sub>152</sub>. En **A** se muestran las interacciones del sitio  $\alpha$ ; en **B**, las del sitio  $\beta$ ; y en **C**, las del sitio  $\gamma$ . Las moléculas de agua se representan mediante *esferas* de color rojo. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno que se encuentran dentro del intervalo 2,0–3,5 Å. Las interacciones se representan en *línea de puntos* y las medidas de las distancias vienen dadas en *ångströms* (Å).

Existen interacciones con el anillo de galactosa que están mediadas por moléculas de agua. En el sitio  $\alpha$  (**Figura 4.30A**), el oxígeno O2 forma tres enlaces de hidrógeno con tres moléculas de agua. Dos de estas moléculas participan en una red de puentes con otras moléculas de agua, mientras que la otra enlaza directamente con el oxígeno del enlace glucosídico. El oxígeno O3 también contribuye a fijar moléculas de agua que participan en la citada red. El oxígeno O4, por su parte, interacciona con uno de los oxígenos O $\delta$  de la cadena lateral del Asp-45. Este residuo, además, reconoce al oxígeno O6, que forma otro enlace de hidrógeno con el nitrógeno N $\epsilon$ -2 de la Gln-46. Asimismo, el residuo de Arg-27 no sólo es

importante en el reconocimiento del anillo de glucosa, sino que forma enlaces de hidrógeno directos con el oxígeno del enlace glicosídico y el oxígeno del anillo de galactosa. Este último oxígeno también une el Nδ-1 de la His-37, aminoácido clave que orienta al Glu-41 para que participe en la red de moléculas de agua conjuntamente con la Tyr-43.

A diferencia de lo que ocurre en el sitio α, en el sitio β (**Figura 4.30B**) existe un menor número de interacciones. El oxígeno O2 únicamente reconoce a una molécula de agua, que también forma parte de un enlace de hidrógeno con el oxígeno del enlace glicosídico, reconocido al mismo tiempo por la Arg-75. El oxígeno O3, por su parte, une dos moléculas de agua, implicadas en una red con otras dos moléculas de agua, que son reconocidas por el residuo de Asp-93. Este aminoácido, además, ancla al anillo de galactosa a través de los oxígenos O4 y O6. Al mismo tiempo, el oxígeno O6 forma un enlace de hidrógeno con el nitrógeno Nδ-2 de la Asn-94.

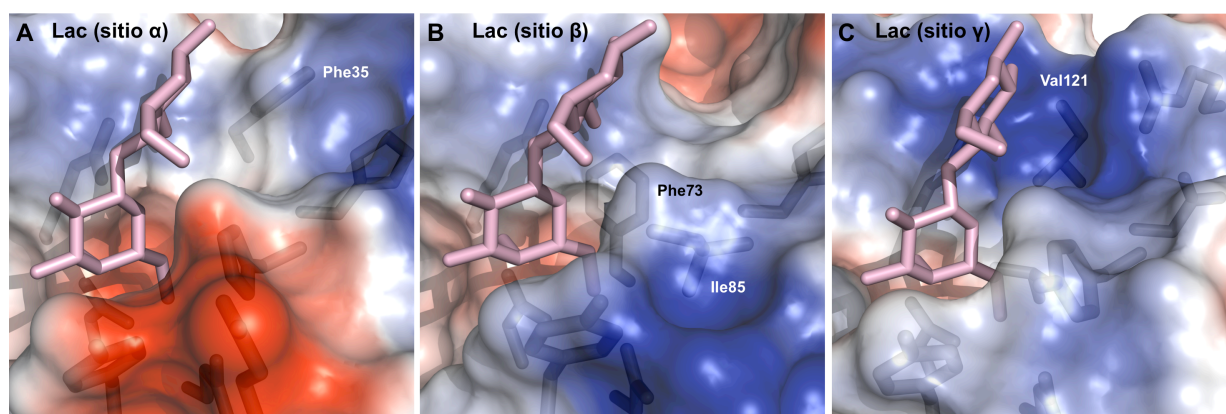
Finalmente, con respecto al sitio γ (**Figura 4.30C**), los oxígenos O2 y O3 solamente interaccionan con moléculas de agua. Una de las moléculas de agua reconocidas por el O2 interviene en un puente con otra molécula de agua, participando ambas en una red en la que intervienen los grupos hidroxilo de los residuos Tyr-42 y Tyr-135 y el oxígeno O6' del anillo de glucosa. El resto de interacciones ya han sido descritas previamente. La Arg-123 reconoce al oxígeno glicosídico y al oxígeno del anillo, el Asp-144 forma enlaces de hidrógeno con los oxígenos O4 y O6, y la Gln-145 interacciona con el oxígeno O6 a través del nitrógeno Nε-2.

Si se tienen en cuenta los contactos que se establecen directamente entre la lactosa y los residuos de los sitios de unión, en el sitio α aparece un mayor número de interacciones (**Tabla 4.5**). Además, es en este sitio donde se identifica un mayor número de residuos implicados en las interacciones mediadas por moléculas de agua. Estos datos podrían apuntar hacia una mayor afinidad del sitio α por el disacárido.

	LBL <sub>152</sub>		
	α	β	γ
<b>Especificidad</b>	β-galactósidos	β-galactósidos	β-galactósidos
<b>Aromático</b>	Tyr43	Tyr91	Tyr142
<b>Ligando-O2</b>	—	—	—
<b>Ligando-O3</b>	—	—	—
<b>Ligando-O4</b>	Asp45	Asp93	Asp144
<b>Ligando-O6</b>	Asp45 y Gln46	Asp93 y Asn94	Asp144 y Gln145
<b>Ligando-O<sub>anillo</sub></b>	Arg27 e His37	—	Arg123

**Tabla 4.5.** Comparación de las interacciones que se establecen en los sitios operativos de unión a azúcar de LBL<sub>152</sub>. Solo se han analizado los residuos que interaccionan directamente con el anillo de galactosa.

Junto a las anteriores consideraciones, cabe mencionar el papel que podrían jugar los residuos apolares del entorno de los sitios de unión. En el sitio  $\beta$ , por ejemplo, están presentes los residuos de Ile-85 y Phe-73 (que interacciona con la Arg-75 a través de una interacción de tipo catión- $\pi$ ), muy próximos al sitio de unión. Tanto el volumen de las cadenas laterales de estos residuos como su carácter hidrofóbico condicionarían la estructura del solvente en el sitio de unión del azúcar. Equivalente a la Phe-73 del sitio  $\beta$ , se encuentran la Phe-35 del sitio  $\alpha$  (que también interacciona con la Arg-27) y la Val-121 del sitio  $\gamma$ . La mayor hidrofobicidad encontrada en el sitio  $\beta$  podría, por tanto, poner en desventaja a este sitio con respecto a la unión en los otros dos sitios disponibles (**Figura 4.31**), es decir, dificultaría el establecimiento de un mayor número de enlaces de hidrógeno. Sin embargo, existe otra lectura en la que los residuos hidrofóbicos se podrían considerar una ventaja en la unión del ligando. Teniendo en cuenta que la presencia del ligando disminuye el área expuesta al solvente en la proteína, debería producirse la liberación de moléculas de agua hacia el solvente durante la unión, lo que podría entenderse como una contribución favorable, puesto que se produce una ganancia de entropía. Además, el retorno de moléculas de agua al medio es un proceso entálpicamente favorable porque se reduce el área expuesta al solvente (desolvatación) y porque se libera energía al romperse los enlaces de hidrógeno formados entre las moléculas de agua estructurales en el estado no unido (Toone, 1994). En conclusión, parece claro que las relaciones entre la estructura atómica y la energética de procesos de unión en sistemas acuosos son ciertamente complejas, más aún en el caso de no disponer de la estructura de la lectina sin el ligando.



**Figura 4.31.** Representación de la superficie de potencial electrostático de LBL<sub>152</sub> sobre los sitios de unión a azúcar. **A.** Se muestra el sitio de unión  $\alpha$ , así como el residuo de Phe35. **B.** Sitio de unión  $\beta$ . Los residuos de Phe73 e Ile85 contribuyen a aumentar la hidrofobicidad. **C.** Se muestra el sitio de unión  $\gamma$ , y el residuo de Val121. La superficie está coloreada de acuerdo al potencial electrostático, azul para la carga positiva y rojo para la negativa.

En último lugar, es destacable la cavidad que aparece en torno al residuo de glucosa en el sitio  $\beta$  (**Figura 4.31B**). Este espacio no se encuentra en los otros sitios de unión, donde aparecen otros residuos involucrados en acomodar mejor a la molécula de lactosa. Este espacio podría emplearse en el reconocimiento de carbohidratos más complejos, es decir, con una cadena más larga que necesite mayor espacio en la superficie. En este sentido, la oquedad podría comportarse como un sitio extendido del sitio de unión  $\beta$ .

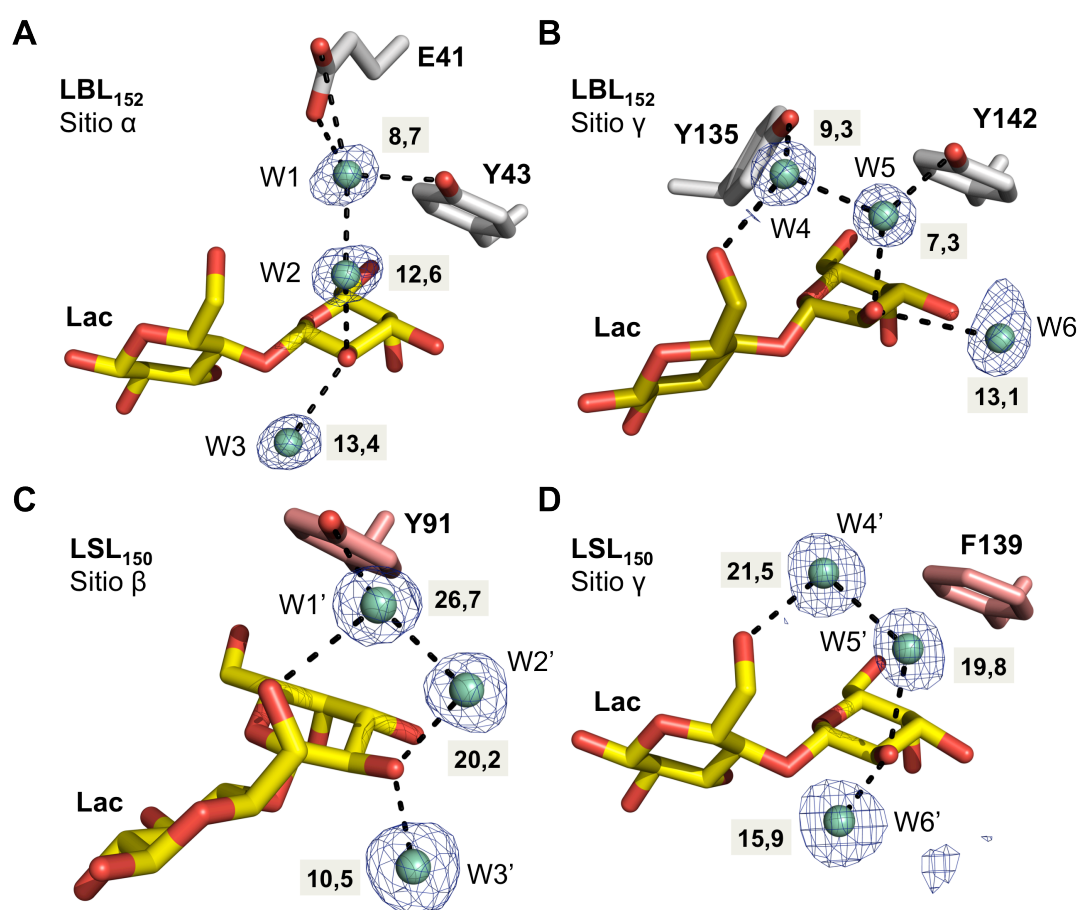
Si se comparan los residuos que intervienen en la unión de lactosa en LBL<sub>152</sub> con los implicados en otros tréboles  $\beta$  independientemente del organismo de origen (**Tablas 4.2 y 4.3**), lo que más llama la atención es la falta de residuos que reconozcan la posición O3 del anillo de galactosa, función atribuida en la mayoría de los casos al ácido aspártico que interacciona también con el oxígeno O4 o minoritariamente a residuos de His y Asn o Gln. LBL<sub>152</sub> suple esta carencia con la intervención de una red de moléculas de agua que trata de fijar la molécula de lactosa, estableciendo enlaces de hidrógeno con las posiciones O2 y O3, al mismo tiempo que con otros aminoácidos que conforman el sitio de unión. Los ejemplos más parecidos al comportamiento del módulo lectina LBL<sub>152</sub> son los que tienen lugar en las lectina MOA del hongo *M. orades* (Grahm *et al.*, 2009) y BEL de *B. edulis* (Bovi *et al.*, 2013). Casualmente, estas lectinas también disponen de los tres sitios potenciales de unión operativos.

#### 2.5.5. Las moléculas de agua son parte de la arquitectura del sitio de unión

Por otro lado, con respecto a las moléculas de agua, se observa cómo éstas juegan un papel fundamental en la unión de la lactosa. Como se ha descrito para otras lectinas y particularmente en el caso de LSL<sub>150</sub>, los enlaces de hidrógeno mediados por moléculas de agua pueden ser tan fuertes como los que se establecen directamente a través de los residuos del sitio de unión (Weis y Drickamer, 1996; Tschampel y Woods, 2003). Aunque no disponemos de la estructura de LBL<sub>152</sub> sin ligando, que permitiría identificar reorganizaciones del solvente como consecuencia de la unión del azúcar, se puede inferir el grado de movilidad de las aguas analizando sus factores *B*. En este sentido, el análisis de estos parámetros ha permitido diferenciar entre aguas con factores *B* cercanos a los de los aminoácidos de su alrededor y aguas con factores *B* más elevados y por lo tanto con mayores desplazamientos con respecto su posición de equilibrio. Si se comparan las aguas de los tres sitios de unión de LBL<sub>152</sub>, se observa que aquéllas con factores *B* más bajos son similares en los sitios  $\alpha$  y  $\gamma$



(Figura 4.32A y B). Estas aguas interaccionan con el oxígeno O2 del anillo de galactosa y con el grupo hidroxilo del residuo de Tyr implicado en el apilamiento del azúcar. En ambos casos, los factores  $B$  tienen un valor medio próximo a  $10 \text{ \AA}^2$ , la mitad aproximadamente que para el resto de moléculas de agua circundantes del sitio de unión. Curiosamente, moléculas de agua en posiciones equivalentes y con bajos factores  $B$  también están presentes en el complejo con lactosa de la lectina LSL<sub>150</sub> (Figura 4.32C y D). Aunque los factores  $B$  de las moléculas de agua en LSL<sub>150</sub> son más elevados, la normalización de estos con respecto al factor  $B$  promedio de toda la proteína demuestra que la movilidad relativa de las moléculas de agua es equivalente a la de LBL<sub>152</sub> (Tabla 4.6). El hecho de que proteínas homólogas evolutivamente conserven moléculas de agua en localizaciones equivalentes en los sitios de unión sugiere que estas moléculas son importantes en la interacción con el azúcar y que se comportan, por tanto, como una extensión del sitio de unión, favoreciendo la correcta acomodación del ligando (Lobsanov *et al.*, 1993).



**Figura 4.33.** Moléculas de agua que ocupan posiciones similares en los sitios de unión de LBL<sub>152</sub> y LSL<sub>150</sub>. En A y B se muestran los sitios de unión  $\alpha$  y  $\gamma$  de LBL<sub>152</sub>, respectivamente; mientras que en C y D, se muestran los sitios  $\beta$  y  $\gamma$  de LSL<sub>150</sub>. Las moléculas de agua se representan como *esferas* de color *cian* junto con el mapa de densidad electrónica  $2F_o - F_c$  (*azul*) a  $1\sigma$  y los valores correspondientes de los factores  $B$  sobre fondo *beige*. Las interacciones en las que están involucradas las moléculas de agua aparecen en *línea de puntos*.

LBL <sub>152</sub>			LSL <sub>150</sub>		
Aguas		$B_{OA} = 7,3 \text{ \AA}^2$	Aguas		$B_{OA} = 14,5 \text{ \AA}^2$
Numeración	Factor $B \text{ (\AA}^2\text{)}$	$K_{\text{agua/lectina}}$	Numeración	Factor $B \text{ (\AA}^2\text{)}$	$K_{\text{agua/lectina}}$
W1	8,7	1,2	W1'	26,7	1,8
W2	12,6	1,7	W2'	20,2	1,4
W3	13,4	1,8	W3'	10,5	0,7
W4	9,3	1,3	W4'	21,5	1,5
W5	7,3	1,0	W5'	19,8	1,4
W6	13,1	1,8	W6'	15,9	1,1

**Tabla 4.6.** Normalización de los factores  $B$  de las moléculas de agua conservadas en los sitios de unión de azúcares de LBL<sub>152</sub> ( $\alpha$  y  $\gamma$ ) y LSL<sub>150</sub> ( $\beta$  y  $\gamma$ ). La relativización de los factores  $B$  se ha llevado a cabo mediante el cálculo de la razón  $K_{\text{agua/lectina}}$  que se obtiene como cociente entre el factor  $B$  asignado a cada una de las moléculas de agua y el factor  $B$  promedio ( $B_{OA}$ ) de cada lectina.

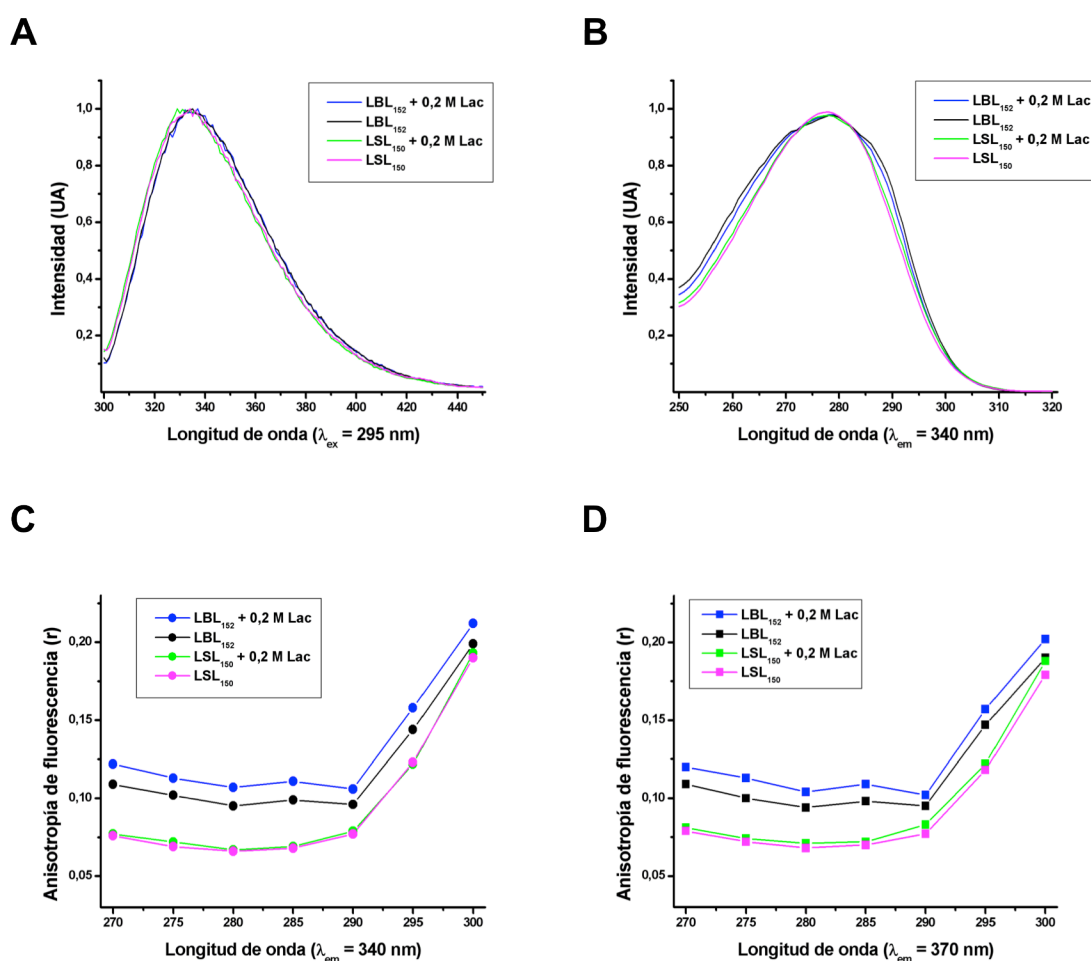
### 2.5.6. La unión de lactosa induce cambios sutiles en la estructura de LBL<sub>152</sub>

Los posibles cambios conformacionales de LBL<sub>152</sub> inducidos por la unión de lactosa se han analizado mediante espectroscopía de fluorescencia en estado estacionario. Estos estudios permiten detectar cambios en la organización tridimensional y en la polaridad del microentorno de los fluoróforos intrínsecos de las proteínas, principalmente, tirosina y triptófano. El sistema LSL<sub>150</sub> se analizó en paralelo como una referencia relativa, teniendo en cuenta que conocemos su estructura tanto en ausencia como en presencia de lactosa.

En primer lugar, se registraron los espectros de emisión de los triptófanos, excitando a una longitud de onda de 295 nm (**Figura 4.33A**). A la vista de los espectros, se observa que no existen diferencias apreciables en presencia o ausencia del ligando. Variaciones en la intensidad de fluorescencia menores del 10% indican que la lactosa no ejerce ningún tipo de desactivación (*quenching*) sobre la capacidad de emisión de los residuos de triptófano, esto es, su rendimiento cuántico no se ve alterado. No obstante, sí se aprecia que el máximo de emisión de los espectros de LBL<sub>152</sub> está desplazado unos 2 nm hacia el rojo comparados con los de LSL<sub>150</sub>. Esto podría indicar que la polaridad de los triptófanos de LBL<sub>152</sub> (Trp-47 y Trp-95) es ligeramente mayor que la de los de LSL<sub>150</sub> (Trp-95 y Trp-131). De hecho, a la vista de las estructuras, se ve cómo los residuos de triptófano de LBL<sub>152</sub> están más expuestos que los de LSL<sub>150</sub>. Por último, cabe señalar que la longitud de onda correspondiente al máximo en los cuatro espectros es alrededor de 335 nm, lo que supone un desplazamiento considerable hacia el azul con respecto al máximo del triptófano en solución (345-350 nm). Esto es consistente con el hecho de que los triptófanos en las estructu-



ras de las lectinas se encuentran parcialmente enterrados y no totalmente expuestos al solvente.

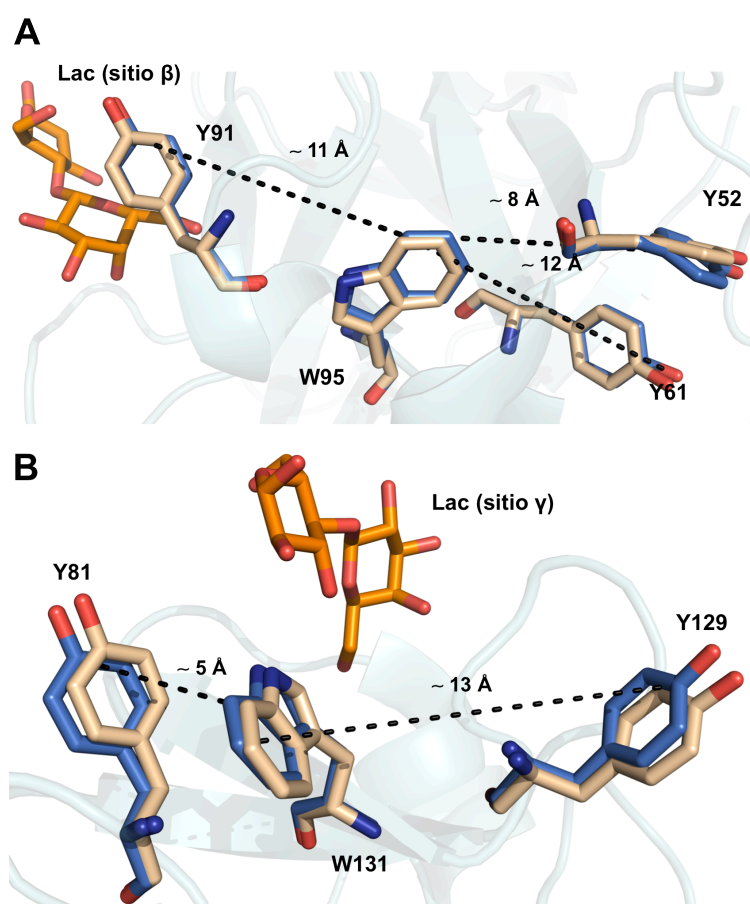


**Figura 4.33.** Espectroscopía de fluorescencia en estado estacionario de LBL<sub>152</sub> y LSL<sub>150</sub> en presencia o no de 0,2 M lactosa. **A.** Espectros de emisión de triptófano corregidos y normalizados. **B.** Espectros de excitación de triptófano corregidos y normalizados. **C y D.** Espectros de anisotropía de excitación de fluorescencia, fijando la longitud de onda de emisión a 340 y 370 nm, respectivamente.

En segundo lugar, se midieron los espectros de excitación de las proteínas, registrando la emisión a 340 nm tras excitación en el intervalo de 250-320 nm (**Figura 4.33B**). A longitudes de onda inferiores a 295 nm no sólo se excitan los triptófanos, sino que también se excitan las tirosinas, que pueden transferir su energía a los triptófanos cercanos, debido al solapamiento del espectro de emisión de la tirosina con el de absorción del triptófano. Este fenómeno puede explicar ciertas diferencias observadas entre los espectros de excitación de LBL<sub>152</sub> y LSL<sub>150</sub>. En el caso de LBL<sub>152</sub> se observa un pequeño aumento en la intensidad en la región de absorción de las tirosinas, y se remarca el hombro en la región de 290 nm, respecto de LSL<sub>150</sub>. Además, en LBL<sub>152</sub> se aprecian pequeñas diferencias en los espectros de excitación cuando hay lactosa en el medio. Esto podría indicar que la lactosa

está modificando la organización relativa de los residuos de tirosina y triptófano en el complejo LBL<sub>152</sub>:lactosa y, por lo tanto, se verían afectados los procesos de transferencia de energía entre ellos.

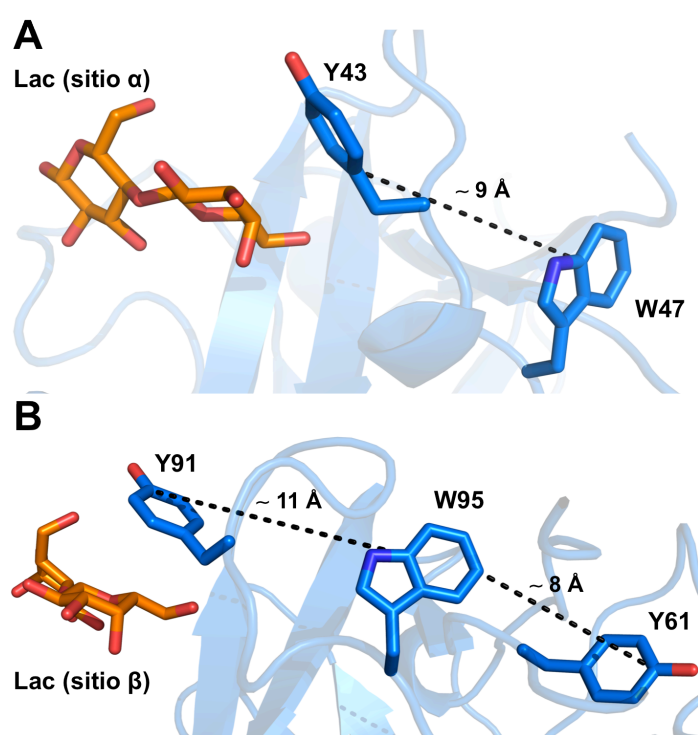
En último lugar, se midió la anisotropía de excitación de fluorescencia ( $r$ ), fijando las longitudes de onda de emisión a 340 y 370 nm (**Figura 4.33C y D**). A estas dos longitudes de onda únicamente emite el triptófano, lo que facilita extraer mayor información acerca de los procesos FRET<sup>36</sup> Tyr→Trp en las lectinas. Un análisis riguroso de los datos de anisotropía implica desglosar estos en dos intervalos diferenciados. El primer intervalo comprende las longitudes de onda de excitación de 270-290 nm, en donde se excitan tanto residuos de tirosina como de triptófano, y que permitiría evaluar cambios en la dinámica de estos dos residuos, además del efecto de la transferencia de energía resonante (FRET) entre ellos (Valeur y Berberan-Santos, 2012). Por otro lado, el segundo intervalo abarca la región de 295-300 nm, donde sólo se excita al triptófano. En este caso, los cambios registrados pueden atribuirse exclusivamente a la dinámica de estos aminoácidos.



**Figura 4.34.** Representación del entorno de los residuos de triptófano de LSL<sub>150</sub> en las estructuras con y sin ligando. **A.** Entorno del Trp-95, localizado en subdominio  $\beta$ . **B.** Entorno del Trp-131, localizado en subdominio  $\gamma$ . Los residuos de Tyr y Trp y las moléculas de lactosa (*naranja*) se muestran con el modelo de *varillas*. Los residuos de la estructura de LSL<sub>150</sub> sin ligando están pintados en *beige*; mientras que los de la estructura con ligando, en *azul*. Las distancias interatómicas se representan mediante *línea de puntos*.

Con respecto a los resultados relativos al primer intervalo de longitudes de onda de excitación (270-290 nm), la anisotropía de fluorescencia indica a grandes rasgos que LSL<sub>150</sub>

presenta valores más bajos que LBL<sub>152</sub>, los cuales se traducen en una mayor eficiencia de FRET Tyr→Trp. A la vista de la estructura y observando en detalle el entorno de los triptófanos de LSL<sub>150</sub> (**Figura 4.34**), se puede justificar este resultado porque existe mayor número de residuos de tirosina alrededor de los residuos Trp-95 y Trp-131, que lo observado en LBL<sub>152</sub>. Además, alguna de estas tirosinas se encuentra a una distancia más corta que las identificadas en LBL<sub>152</sub> (**Figura 4.35**). También hay que destacar que la anisotropía no cambia en LSL<sub>150</sub> cuando la lactosa está presente en el medio, lo que está de acuerdo con lo observado en la estructura cristalográfica. Ambas estructuras, con y sin ligando, son equivalentes casi por completo, por lo que la eficiencia FRET no se ve afectada.



**Figura 4.35.** Representación del entorno de los residuos de triptófano de LBL<sub>152</sub> en las estructura del complejo con lactosa. **A.** Entorno del Trp-47, localizado en subdominio  $\alpha$ . **B.** Entorno del Trp-95, localizado en subdominio  $\beta$ . Los residuos de Tyr y Trp (azul) y las moléculas de lactosa (naranja) se muestran con el modelo de varillas. Las distancias interatómicas se representan mediante línea de puntos.

Por otro lado, el aumento en un 45% aproximadamente del valor de la anisotropía de LBL<sub>152</sub> con respecto a la de LSL<sub>150</sub> indica que los fenómenos de FRET están siendo mucho menos eficientes (Weber, 1960; Valeur y Weber, 1977). Si a esto se le suma el hecho de que LBL<sub>152</sub> experimenta un incremento del 10% de la anisotropía cuando la lactosa está unida, todo indica que existe un efecto en la estructura inducido por la lactosa. Este resultado es muy interesante desde un punto de vista de la dinámica del módulo lectina. Teniendo en cuenta que la eficiencia del FRET Tyr→Trp varía con la sexta potencia de la distancia Tyr–Trp, la conclusión principal indicaría que la unión de lactosa promueve un pequeño cambio en la conformación (la distancia o distancias Tyr–Trp y/o sus orientaciones relativas aumentan), lo que conduciría a una menor eficiencia FRET promedio. Esta hipóte-

sis se justifica al observar que en la estructura cristalográfica de LBL<sub>152</sub> (**Figura 4.32**) los residuos de tirosina Tyr-43 y Tyr-91, que están muy próximos a los respectivos aminoácidos Trp-47 y Trp-95, responsables de la emisión de fluorescencia, juegan un papel clave en el reconocimiento de la lactosa.

En el segundo intervalo (295-300 nm) no se observan cambios en la anisotropía de fluorescencia de excitación de LSL<sub>150</sub> con y sin lactosa. Este resultado vuelve a estar de acuerdo con que el microentorno de los triptófanos en este módulo lectina no cambia esencialmente al formarse el complejo. Finalmente, el hecho de que en el caso de LBL<sub>152</sub> se observe un incremento constante de aproximadamente el 10% de la anisotropía de fluorescencia en todo el intervalo de longitudes de onda de excitación (270-300 nm) parece indicar que al formarse el complejo de LBL<sub>152</sub> con lactosa, la movilidad de uno o los dos triptófanos estaría más restringida en el complejo.

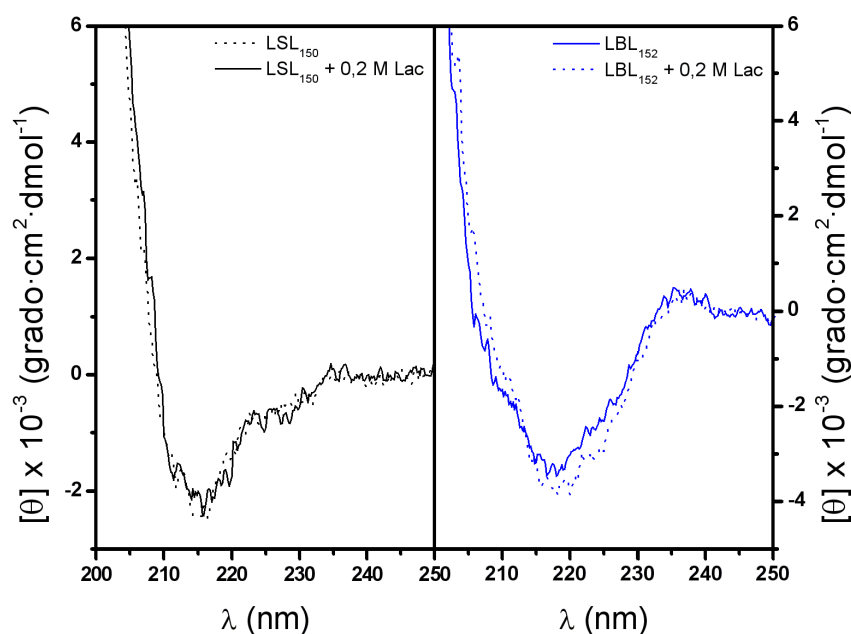
El conjunto de resultados de los experimentos de fluorescencia revelan que la estructura global de LBL<sub>152</sub> sin ligando debería ser prácticamente idéntica a la obtenida en complejo con lactosa mediante cristalografía. Esto es consistente con lo observado para las lectinas en general, puesto que la unión de los ligandos no suele implicar un gran cambio conformacional evidente (Gohler *et al.*, 2010; Gabius *et al.*, 2011). Sin embargo, sí se observan cambios a nivel local que indicarían que el módulo lectina de *L. bicolor* es menos compacto en el sentido de que presenta una estructura más flexible en ausencia de ligando. Por esta razón, la unión de lactosa produciría el aumento observado en la anisotropía de fluorescencia, reflejando una pérdida de esta libertad de movimiento, al menos, en el entorno de los triptófanos.

#### 2.5.7. La estructura global de LBL<sub>152</sub> no sufre cambios por la unión del ligando

Como complemento a los experimentos de cristalografía y fluorescencia, se llevaron a cabo medidas de dicroísmo circular (DC) en la región ultravioleta con objeto de analizar potenciales cambios en la estructura. Estos espectros son útiles porque aportan información tanto de la estructura secundaria (UV lejano) como de la estructura terciaria (UV cercano) de las proteínas en solución (Kelly *et al.*, 2005). Los ensayos de dicroísmo circular se llevaron a cabo con los módulos lectina LBL<sub>152</sub> y LSL<sub>150</sub>.

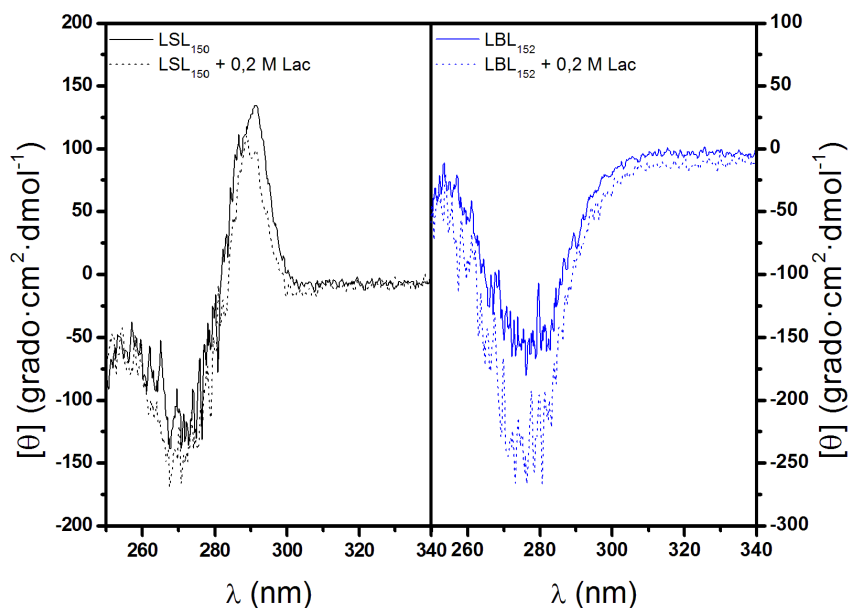
Los espectros de DC en el UV lejano de las dos lectinas son similares (**Figura 4.36**), lo cual es reflejo de su parecida estructura secundaria. En ambos casos se observa un mí-

nimo en torno a los 215-218 nm, que refleja el alto contenido en estructura secundaria de tipo lámina  $\beta$ . El espectro de LBL<sub>152</sub> tiene, además, un pequeño máximo centrado cerca de los 235 nm, que también aparece en otros tréboles  $\beta$  como, por ejemplo, en la lectina CNL de *C. nebularis* (Pohleven *et al.*, 2012), el cual se asocia con una señal procedente de residuos de triptófano que se encuentran en un entorno singular debido al plegamiento de la proteína (Woody, 1994). Por otro lado, no se observa ningún cambio sustancial por efecto de la lactosa, ya que los espectros coinciden casi perfectamente.



**Figura 4.36.** Espectroscopía de dicroísmo circular en el UV lejano. Comparación de los espectros de dicroísmo circular de LSL<sub>150</sub> (negro) y LBL<sub>152</sub> (azul) en ausencia (línea continua) y presencia (línea discontinua) de 0,2 M lactosa. Los espectros en presencia de ligando se obtuvieron dializando previamente las lectinas con tampón TSAL. Los espectros fueron medidos a 25 °C y con una cubeta de 0,1 cm de paso óptico.

A pesar de que la estructura tridimensional es similar, los espectros de DC en el UV cercano son totalmente diferentes para LSL<sub>150</sub> y para LBL<sub>152</sub> (Figura 4.37), lo cual no es extraño en tanto que estos espectros resultan de aspectos sutiles de la estructura terciaria de las proteínas, sirviendo de huella dactilar específica de cada proteína. Los principales cromóforos en esta región del espectro y, por ello, los principales responsables de la actividad óptica son los residuos aromáticos. Sorprendentemente, y en consonancia con lo observado mediante los experimentos de anisotropía de excitación de fluorescencia, el espectro de LBL<sub>152</sub> cambia cuando hay lactosa, en el medio en tanto que el de LSL<sub>150</sub> se mantiene prácticamente invariable.

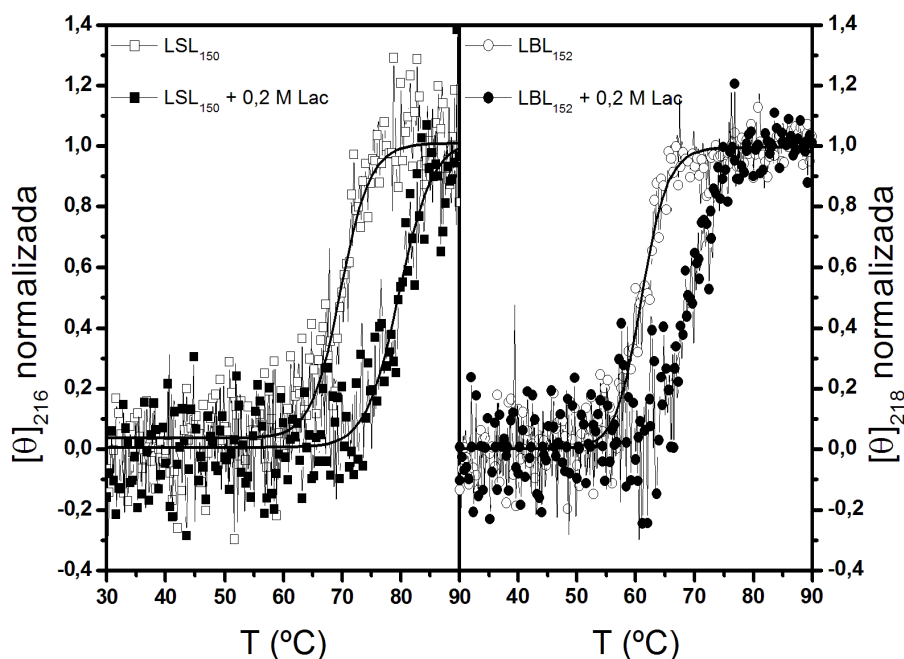


**Figura 4.37.** Espectroscopía de dicroísmo circular en el UV cercano. Comparación de los espectros de dicroísmo circular de LSL<sub>150</sub> (*negro*) y LBL<sub>152</sub> (*azul*) en ausencia (*línea continua*) y presencia (*línea discontinua*) de 0,2 M lactosa. Los espectros en presencia de ligando se obtuvieron dializando previamente las lectinas con tampón TSAL. Los espectros fueron medidos a 25 °C y con una cubeta de 1 cm de paso óptico.

## 2.6. La lactosa ejerce un efecto estabilizante sobre los tréboles β

La caracterización biofísica del módulo lectina N-terminal de *L. bicolor* se completó con el estudio de su estabilidad térmica. En primer lugar, se registró la variación de la elipticidad molar a 218 nm (mínimo del espectro de DC en el UV lejano) en función de la temperatura, que se fue incrementando a una velocidad de 40 °C/h (**Figura 4.38**). El ajuste de los datos experimentales con una función gaussiana/sigmoidea (**Ecuación 4.2** de *Materiales y Métodos*) permitió determinar la temperatura de desnaturalización ( $T_m$ ), que para LBL<sub>152</sub> fue de  $61,0 \pm 0,2$  °C. Este valor es inferior al obtenido para el módulo lectina LSL<sub>150</sub> ( $69,4 \pm 0,3$  °C), cuya  $T_m$  se obtuvo igualmente registrando la elipticidad molar, pero usando una longitud de onda de 216 nm. La desnaturalización térmica provocó en ambos casos la precipitación masiva de la proteína desnaturalizada, lo que implica que su desnaturalización es irreversible.

Un experimento similar se realizó incubando previamente la proteína con 0,2 M lactosa (**Figura 4.38**). Para ambas lectinas, se observa un efecto estabilizante registrándose una  $T_m$  de  $69,4 \pm 0,3$  °C para LBL<sub>152</sub> y  $81,2 \pm 0,3$  °C para LSL<sub>150</sub>. Este efecto estabilizante de la lactosa ya había sido descrito previamente para otras lectinas, como, por ejemplo, miembros de la familia de las galectinas (Ermakova *et al.*, 2013).



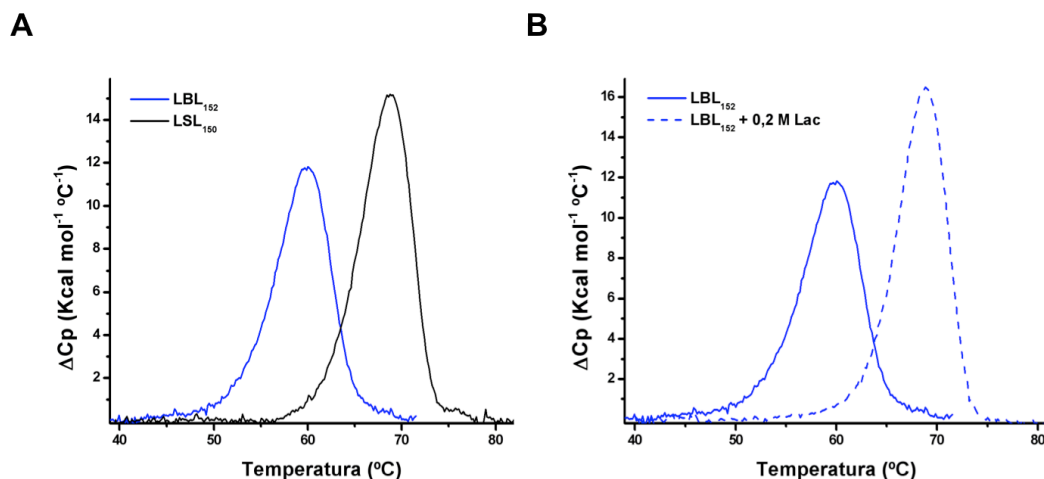
**Figura 4.38.** Efecto de la lactosa en la estabilidad térmica de los módulos lectina N-terminales LSL<sub>150</sub> y LBL<sub>152</sub>. La elipticidad a 216 nm para LSL<sub>150</sub> (cuadrados) y a 218 nm para LBL<sub>152</sub> (círculos) fue monitorizada como una función de la temperatura en ausencia (símbolo vacío) y presencia (símbolo lleno) de 0,2 M lactosa. La línea continua se corresponde al ajuste a una curva sigmoidea de los datos experimentales para la determinación de la temperatura media de transición ( $T_m$ ).

Asimismo, se han llevado a cabo ensayos de calorimetría diferencial de barrido (DSC) (**Figura 4.39**). De esta manera, se determinó una  $T_m$  de  $60,03 \pm 0,03$  °C para LBL<sub>152</sub> y una de  $68,9 \pm 0,02$  °C para el correspondiente complejo con lactosa, lo que implica una  $\Delta T_m$  de 9 °C, resultado prácticamente idéntico al obtenido por DC (**Figura 4.39B**). De igual forma, el comportamiento en el DSC de LSL<sub>150</sub> resultó igual al obtenido mediante DC, determinándose una  $T_m$  de  $68,77 \pm 0,02$  °C para la lectina sin lactosa (**Figura 4.39A**). La integración de las curvas obtenidas mediante DSC permitió obtener un valor de la entalpía de desnaturalización, siendo de  $100 \pm 10$  kcal/mol y  $102 \pm 2$  kcal/mol para LBL<sub>152</sub> en ausencia y presencia de lactosa, respectivamente. La lactosa, por tanto, no afecta a la energética del desplegamiento. La entalpía determinada para LSL<sub>150</sub> está asimismo dentro del mismo rango ( $105 \pm 5$  kcal/mol). El valor de las entalpías indica que se trata de un proceso endotérmico y la forma del pico advierte que la desnaturalización se produce cooperativamente.

Las lectinas fúngicas y, en concreto, los tréboles  $\beta$  caracterizados suelen tener  $T_m$  elevadas. Este comportamiento podría indicar una relación entre simetría interna y estabilidad y, de hecho, esta relación ha sido demostrada experimentalmente mediante ingeniería de proteínas, utilizando tréboles  $\beta$  (Lee y Blaber, 2011; Broom *et al.*, 2012). Por otro lado, se ha visto que la incorporación de sitios funcionales de unión puede afectar negativamente a la estabilidad (Shoichet *et al.*, 1995). Este hecho podría justificar de algún modo



por qué LBL<sub>152</sub> es en principio más inestable que LSL<sub>150</sub>, ya que presenta un sitio adicional para unir lactosa. Parece que la presencia de sitios funcionales en proteínas no implica evolutivamente las soluciones más estables sino una situación de compromiso entre dinámica y arquitectura asociadas a la función y pura estabilidad termodinámica (Gosavi *et al.*, 2008).



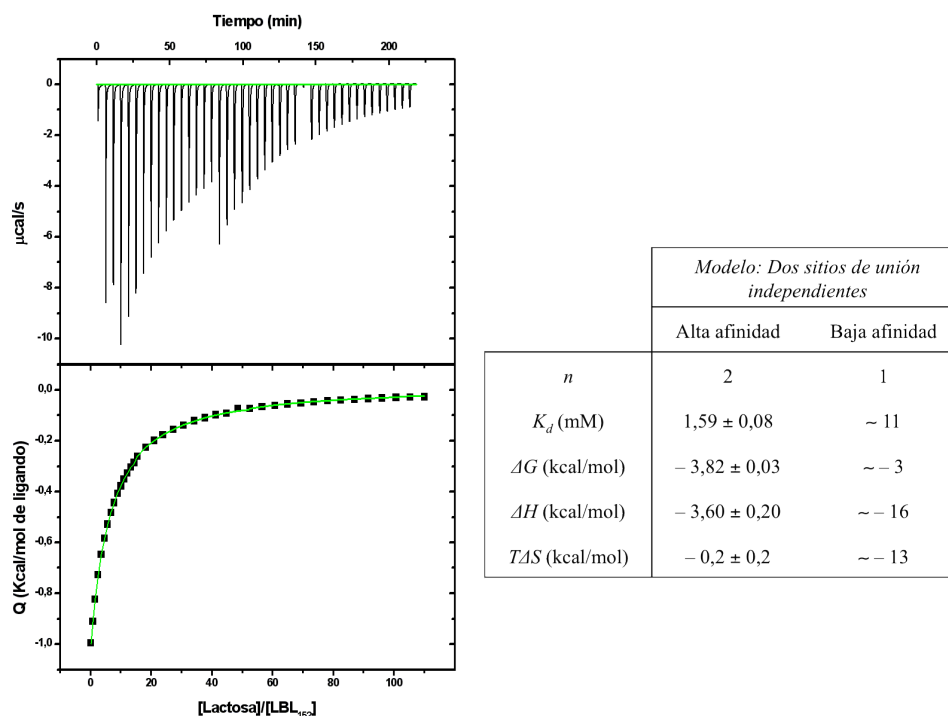
**Figura 4.39.** Estudio mediante calorimetría diferencial de barrido (DSC) del efecto de la lactosa en la estabilidad térmica de los módulos lectina N-terminales LSL<sub>150</sub> (*negro*) y LBL<sub>152</sub> (*azul*). La capacidad calorífica fue monitorizada como una función de la temperatura en ausencia (*línea continua*) y presencia (*línea discontinua*) de 0,2 M lactosa.

## 2.7. Caracterización termodinámica de la interacción de LBL<sub>152</sub> con lactosa

La afinidad por lactosa del módulo lectina N-terminal de *L. bicolor* fue estudiada mediante ITC. Los experimentos de ITC se llevaron a cabo a 25 °C en tampón TSA utilizando lactosa a una concentración final de 60 mM. El perfil de unión del ligando (**Figura 4.40**) demuestra que la unión es exotérmica y la curva de titulación se ajusta perfectamente a un modelo de dos tipos de sitios de unión independientes.

Debido a la baja afinidad que el módulo lectina presenta por la lactosa (no se aprecia saturación en la curva de valoración), la estequiometría se tuvo que ajustar manualmente. Este problema es frecuente en muchas lectinas, ya que éstas se caracterizan por poseer constantes de afinidad muy bajas para la unión de mono y disacáridos, en el rango milimolar (Gabiús, 2009). Este rango no es el adecuado para el cálculo preciso de los correspondientes parámetros termodinámicos. Para ello, sería necesario que se cumpliera la relación  $[LBL_{152}] \cdot n \cdot K_d^{-1} \geq 1 - 5$  (Wiseman *et al.*, 1989). Esta consideración supondría usar una concentración de proteína unas quince veces superior a la utilizada en este experimento (~4 mg/mL), lo que no es factible desde un punto de vista metodológico.



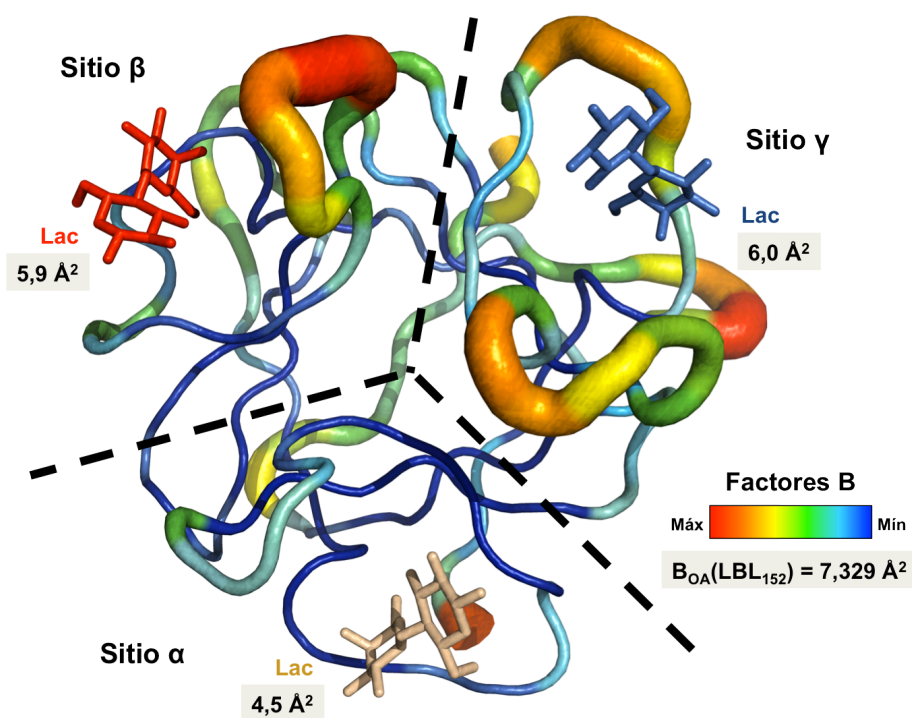


**Figura 4.40.** Análisis mediante ITC de la unión en solución de lactosa a LBL<sub>152</sub>. A la izquierda se muestra la curva de valoración, así como el ajuste de los datos (verde) con el modelo de dos sitios de unión independientes. A la derecha se muestra una tabla con el valor de los parámetros termodinámicos determinados: la estequiometría ( $n$ ), la constante de disociación ( $K_d$ ), la energía de Gibbs ( $\Delta G$ ), la entalpía de unión ( $\Delta H$ ) y el cambio de entropía ( $T\Delta S$ ).

Aún así, la técnica de ITC permitió estimar de forma aproximada la afinidad de la lectina por el azúcar lactosa en solución. En primer lugar, se identificaron dos sitios independientes de unión con una afinidad más alta, siendo la constante de disociación ( $K_{d1}$ ) de  $1,59 \pm 0,08$  mM ( $\Delta G_1 = -3,82 \pm 0,03$  kcal/mol) con una entalpía ( $\Delta H_1$ ) de  $-3,6 \pm 0,2$  kcal/mol. Por otro lado, habría un solo sitio de baja afinidad para el que se obtuvo una constante de disociación ( $K_{d2}$ ) de  $\sim 11$  mM ( $\Delta G_2 = \sim -3$  kcal/mol) y una entalpía ( $\Delta H_2$ ) de  $\sim -16$  kcal/mol. Los valores de entalpía en ambos casos indican que la interacción con la lactosa está dirigida entálpicamente, lo cual es reflejo de la fuerza de las interacciones con la proteína (formación de enlaces de hidrógeno e interacciones de tipo van der Waals), relativa a las existentes previamente con el solvente. El ajuste de la estequiometría se basó en los resultados cristalográficos previos, que demostraron la existencia de tres sitios de unión por molécula de lectina. A diferencia de lo que ocurría en el caso de LSL<sub>150</sub>, los datos experimentales en este caso no pudieron ajustarse a un modelo de dos o más sitios de unión secuenciales, lo que descarta *a priori* la existencia de una secuencia ordenada en la unión del ligando.

La comparación de estos resultados con los obtenidos para LSL<sub>150</sub> indica que, aunque LBL<sub>152</sub> tiene un sitio de unión operativo más, la afinidad es más baja en general. En

términos absolutos, LBL<sub>152</sub> es capaz de reclutar más moléculas de ligando, pero su unión es peor que la que tiene lugar en LSL<sub>150</sub>. A la luz de las estructuras es complicado explicar esta diferencia en la afinidad, puesto que la naturaleza de los contactos que se establecen y los residuos que intervienen en los sitios de unión son parecidos en ambas lectinas. En este caso, debe de ser crítica la contribución de otros factores, además del número de contactos que se establezcan, como la disminución del área expuesta al solvente, el efecto entrópico de las moléculas de agua, la desolvatación del ligando, etc. Estos resultados no se pueden extrapolar para la unión de otros ligandos o de su diana biológica. Generalmente, los azúcares que reconocen las lectinas en el medio natural son más complejos y utilizan otras zonas de reconocimiento adicional, conocidas como sitios extendidos o subsitios, que se disponen en el entorno de los sitios de unión, repercutiendo en un aumento considerable de la afinidad (Weis y Drickamer, 1996).



**Figura 4.41.** Representación de la cadena principal de LBL<sub>152</sub> en función de los factores *B* de cada residuo. Cada uno de los sitios de unión se representa junto con las moléculas de lactosa en modelo de varillas y coloreadas en beige (sitio α), rojo (sitio β) y azul (sitio γ). Los factores *B* se distribuyen en función del intervalo que va de azul a rojo, donde el azul se corresponde a los valores más bajos y el rojo, a los más altos dentro de la estructura.

La interpretación de los resultados de ITC ha de hacerse necesariamente teniendo en cuenta la estructura cristalográfica de LBL<sub>152</sub>. En este sentido, hemos comprobado que el mejor ajuste se consigue considerando tres sitios de unión independientes, uno de menor afinidad y otros dos de mayor afinidad por lactosa. Teniendo en cuenta tanto este resultado como el estudio profundo que se llevó a cabo de las interacciones lectina-azúcar en cada sitio de unión, sugerimos que los sitios α y γ serían probablemente los de mayor afinidad, mientras que el β sería el que poseería una peor capacidad para unir lactosa, al tener

menor número de contactos entre la lectina y el azúcar. Sin embargo, considerando no sólo el número de interacciones, sino también la flexibilidad de la proteína, reflejada en el valor de los factores  $B$ , no se puede asegurar con rotundidad que los sitios  $\alpha$  y  $\gamma$  y sean los de mayor afinidad. La **Figura 4.41** muestra que el sitio de unión más rígido es el sitio  $\alpha$ , por lo que en este subdominio la unión del ligando está favorecida y es más estable. Los subdominios  $\beta$  y  $\gamma$ , en cambio, poseen unos bucles en torno al sitio de unión cuya movilidad es más elevada, lo que podría contribuir a la desestabilización de la unión. De hecho, los factores  $B$  asociados a las moléculas de lactosa van en el mismo sentido; mientras que la lactosa del sitio  $\alpha$  posee un factor  $B$  de tan solo  $4,5 \text{ \AA}^2$ , las otras dos moléculas de lactosa tienen factores  $B$  más altos ( $5,9$  y  $6,0 \text{ \AA}^2$  para los sitios  $\beta$  y  $\gamma$ , respectivamente). Por lo tanto, considerando el análisis de la estructura cristalográfica, solamente podríamos asegurar con mayor probabilidad que el sitio  $\alpha$  es un sitio de alta afinidad comparado con los otros dos existentes.

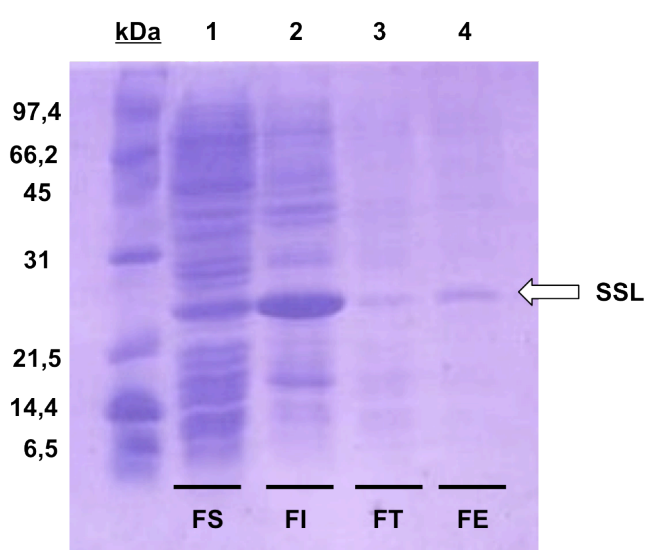
### 3. La lectina SSL de *Sclerotinia sclerotiorum*

#### 3.1. La lectina SSL de *S. sclerotiorum* forma agregados solubles

Gracias a la secuenciación del genoma del hongo ascomiceto *Sclerotinia sclerotiorum* (Amselem *et al.*, 2011), se pudo identificar un nuevo marco de lectura (identificador génico: 4488290) que codificaría una lectina con un 34% de similitud de secuencia con LSLa (Uniprot: A7EMH9\_SCLS1). Con el objetivo de caracterizar esta proteína, en primer lugar el gen fue sintetizado y clonado en el vector pET-28a(+). Células de *E. coli* BL21(DE3) fueron transformadas con el plásmido resultante, donde se llevó a cabo la sobreexpresión de la proteína. El protocolo estándar de purificación de lectinas, que se había aplicado con anterioridad en los casos de LSL<sub>150</sub> y LBL<sub>152</sub>, fue aplicado para intentar aislar la hipotética lectina de *S. sclerotiorum* (SSL). Sin embargo, aunque a la vista del gel (**Figura 4.42**) se observa que la lectina se produjo con un gran rendimiento, la fracción purificada es apenas perceptible, lo cual revela serios problemas de solubilidad.

Con la intención de mejorar el protocolo de purificación, se probaron diferentes estrategias: adición de agentes reductores (4mM de  $\beta$ -mercaptoetanol), adición de detergentes (0,1% de Triton X-100), uso de diferentes concentraciones de sal en el tampón de lisis (10-200 mM de NaCl), utilización de distintas resinas de afinidad (asialofetuína y lactosil-Sepharose®), aplicación de diferentes métodos cromatográficos (intercambio iónico y cro-

matografía de hidrofobicidad) e incluso la preparación de una construcción nueva utilizando His-LSLt (ver la sección 7.1 de *Materiales y Métodos*) como etiqueta de fusión en el extremo N-terminal. En ninguno de estos casos se consiguió obtener una mejora del rendimiento de la purificación de SSL. Teniendo en cuenta que en la fracción insoluble del lisado celular (**Figura 4.42**) había gran cantidad de proteína, probablemente en forma de cuerpos de inclusión, se probó también a desnaturalizar estos agregados mediante el uso de agentes caotrópicos, como 8 M urea. En este caso, SSL precipitaba masivamente cuando la urea se retiraba de la solución, usando diferentes estrategias como, por ejemplo, la dilución infinita o la diálisis exhaustiva.



**Figura 4.42.** SDS-PAGE (15% de poliacrilamida) de la purificación de SSL mediante cromatografía de afinidad a Sepharose® 4B. Carril 1: fracción soluble (FS) del extracto celular. Carril 2: fracción insoluble (FI) del lisado celular. Carril 3: fracción no retenida o *flow-through* (FT). Carril 4: fracción obtenida mediante elución con 0,2 M lactosa.

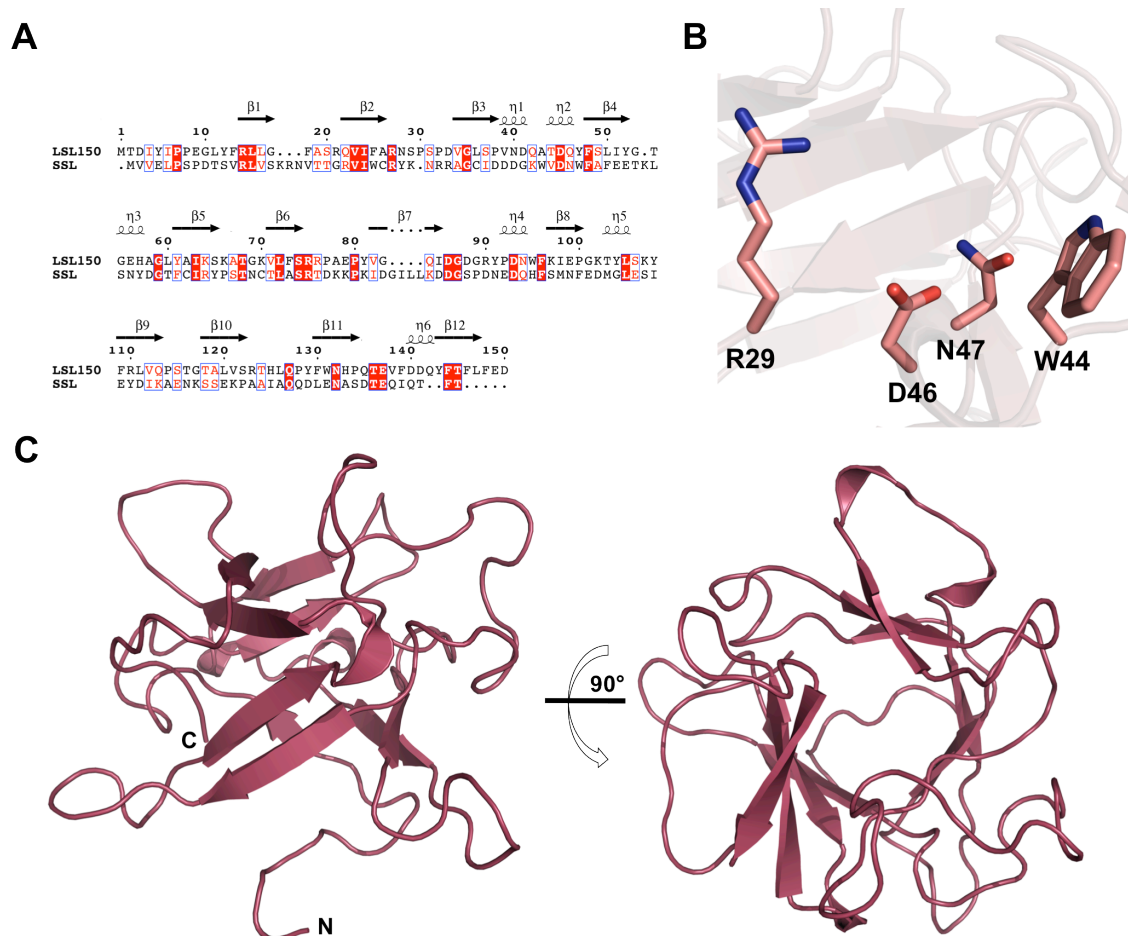
El análisis mediante cromatografía de exclusión por tamaño indicó que tanto SSL como la versión con la etiqueta de fusión en el N-terminal (His-LSLt-SSL) tienden a la formación de agregados solubles que eluían en el volumen de exclusión (datos no mostrados). La agregación fue imposible de revertir, lo cual impidió el estudio funcional y estructural de la correspondiente lectina. Desafortunadamente, este comportamiento hizo que SSL no superase el criterio de suficiencia establecido en los objetivos para validar a estas proteínas como lectinas potenciales para futuras aplicaciones biotecnológicas. Este hecho nos llevó a no continuar con el estudio de esta lectina.

Tal y como se indicó en la *Introducción*, en el análisis llevado a cabo por Kellens y colaboradores, donde se estudiaban las lectinas aisladas directamente del hongo, no se identificó ninguna lectina que coincidiera con SSL (Kellens *et al.*, 1992). De hecho, la única lectina caracterizada de *S. sclerotiorum* hasta la fecha es SSA, cuya estructura es un trébol  $\beta$  dimerico (Sulzenbacher *et al.*, 2010). No obstante, estos resultados no descartan que SSL se

pueda expresar en el hongo bajo determinadas condiciones fisiológicas, lo que dificultaría su aislamiento y caracterización.

### 3.2. SSL sería un nuevo trébol $\beta$ con un único sitio de unión a azúcares

El alineamiento de las secuencias de LSL<sub>150</sub> (homólogo con mayor similitud de secuencia) y SSL revela que este último podría adquirir un plegamiento tipo trébol  $\beta$  (**Figura 4.43A**). Uno de los factores más relevantes es que se puede identificar a lo largo de la secuencia el triplete (QXW)<sub>3</sub> típico de este tipo de estructuras (Hazes, 1996): NWF (subdominio  $\alpha$ ), QHF (subdominio  $\beta$ ) y QTF (subdominio  $\gamma$ ). Como ocurre tanto con LSL<sub>150</sub> como con LBL<sub>152</sub>, el Trp canónico se sustituiría por un residuo de Phe en los tres casos. SSL sería además una lectina trébol  $\beta$  de dominio único, esto es, no poseería ningún dominio C-terminal acompañante, puesto que la secuencia completa son 149 aminoácidos, longitud aproximada de la mayoría de tréboles  $\beta$  ya descritos.



**Figura 4.43.** Estudio *in silico* de la estructura de SSL. **A.** Alineamiento de las secuencias de LSL<sub>150</sub> y SSL en el que se muestran los elementos de estructura secundaria (hélices ( $\eta$ ) y hebras  $\beta$ )). **B.** Disposición de los aminoácidos que intervendrían en la unión del ligando en el sitio  $\alpha$ . **C.** Vistas frontal (izquierda) y cenital (derecha) del modelo obtenido para la estructura de SSL. El modelo se ha obtenido mediante el alineamiento de la secuencia de SSL con la de LSL<sub>150</sub> (ClustalW2, EMBL-EBI) y posterior predicción con el servidor SWISS-MODEL ([www.swissmodel.expasy.org](http://www.swissmodel.expasy.org)).

Utilizando como elemento de partida el alineamiento de secuencias entre LSL<sub>150</sub> y SSL, realizado con el servidor ClustalW2 (EMBL-EBI) (Larkin *et al.*, 2007) y teniendo en cuenta que la estructura de LSL<sub>150</sub> es conocida, se llevó a cabo la modelización de la estructura terciaria de SSL mediante la herramienta de modelado Alignment Mode del servidor SWISS-MODEL (Arnold *et al.*, 2006) (**Figuras 4.43C**). La zona mejor modelada se corresponde con los primeros ~70 residuos, pertenecientes al subdominio  $\alpha$ . Concretamente, en este subdominio estaría el único sitio de unión a azúcares de la lectina, dado que conserva el residuo aromático (Trp-44) responsable de la interacción de apilamiento con el ligando, así como el resto de residuos con probabilidad de establecer enlaces de hidrógeno (**Figura 4.43B**). Este comportamiento es similar al observado en la lectina SSA de *S. sclerotiorum*, cuyo único sitio operativo es el sitio  $\alpha$  (Sulzenbacher *et al.*, 2010).

Finalmente, SSL tiene cuatro cisteínas en la secuencia. Curiosamente, el modelo obtenido por el servidor SWISS-MODEL predice que dos de ellas (Cys-28 y Cys-72) estarían lo suficientemente cerca en la estructura como para formar un puente disulfuro que estabilizase la conformación nativa. Este puente disulfuro podría ser una de las principales razones que dificultan la producción con un buen rendimiento de la lectina SSL en las células de *E. coli* BL21(DE3).

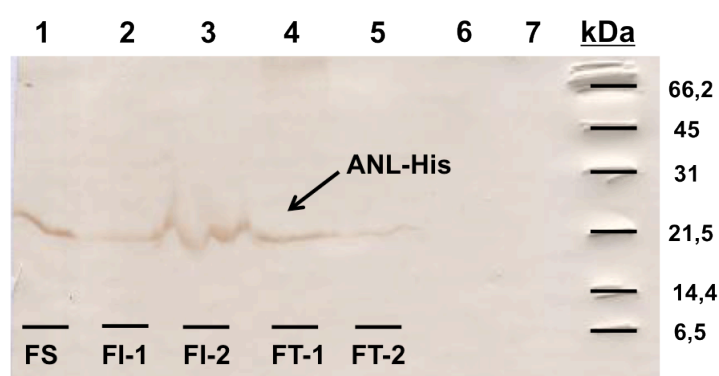
## 4. La lectina ANL de *Aspergillus niger*

### 4.1. La lectina ANL de *A. niger* se obtiene con un rendimiento muy bajo

El último homólogo de LSL<sub>150</sub> que hemos identificado mediante comparaciones de secuencias fue la proteína A2R2Y9\_ASPNC (Uniprot) de *Aspergillus niger* (ANL). La secuencia del genoma de este hongo se reportó en el año 2007 (Pel *et al.*, 2007). Nuestros análisis muestran que esta proteína sería una lectina con un 28% de identidad de secuencia con el módulo lectina N-terminal LSL<sub>150</sub>, y estaría constituida por un solo dominio que adoptaría un plegamiento tipo trébol  $\beta$ .

Con la intención de producir la lectina ANL de *A. niger*, el gen *anl* fue sintetizado y clonado en el vector pURI3-Cter (Curiel *et al.*, 2011), introduciéndose de esta forma una cola de seis histidinas en posición C-terminal (ANL-His). La estrategia de clonación en este caso fue distinta a la utilizada en las otras lectinas porque la primera tentativa de expresión con la construcción en el vector pET-28a(+) no rindió ningún producto soluble tras la cromatografía de afinidad a Sepharose® 4B. De este modo, el gen *anl* fue subclonado en el vector pURI3-Cter, que permite incorporar una etiqueta de afinidad para usar en la purifica-

ción otro método cromatográfico alternativo. La expresión se llevó a cabo en células de *E. coli* BL21(DE3) y la purificación se realizó mediante cromatografía de afinidad a níquel ( $\text{Ni}^{2+}$ ). La proteína no se purificó con éxito, aunque se observó mediante *western blot* (anticuerpos anti-His) que ésta se producía, pero con muy bajo rendimiento (**Figura 4.44**). La banda más intensa aparecía en la fracción insoluble del lisado celular, lo que indicaba que probablemente ANL-His no conseguía plegarse adecuadamente en el interior celular. Con la intención de recuperar la proteína de la fracción insoluble, ésta se trató con 8 M urea. La estrategia de desnaturalización no funcionó pues la proteína precipitaba cuando se eliminaba el agente caotrópico. Por otro lado, la coexpresión con el sistema de chaperonas de *E. coli* GroEL/GroES condujo a la obtención de ANL-His contaminada con GroEL, resultado similar al ocurrido con LBL. La eliminación de GroEL mediante la disociación del complejo tampoco fue satisfactoria.



**Figura 4.44.** Análisis mediante *western blot* de la producción de ANL-His. Se muestra la fracción soluble del extracto celular (FS), así como las fracciones insolubles antes (FI-1) y después de lavar con 8 M urea (FI-2). También se muestran las fracciones no retenidas que eluyen de la columna antes (FT-1) y después del tratamiento con 8M urea (FT-2). La inmunodetección se llevó a cabo usando como anticuerpo primario un anticuerpo anti-His de conejo (Sigma-Aldrich).

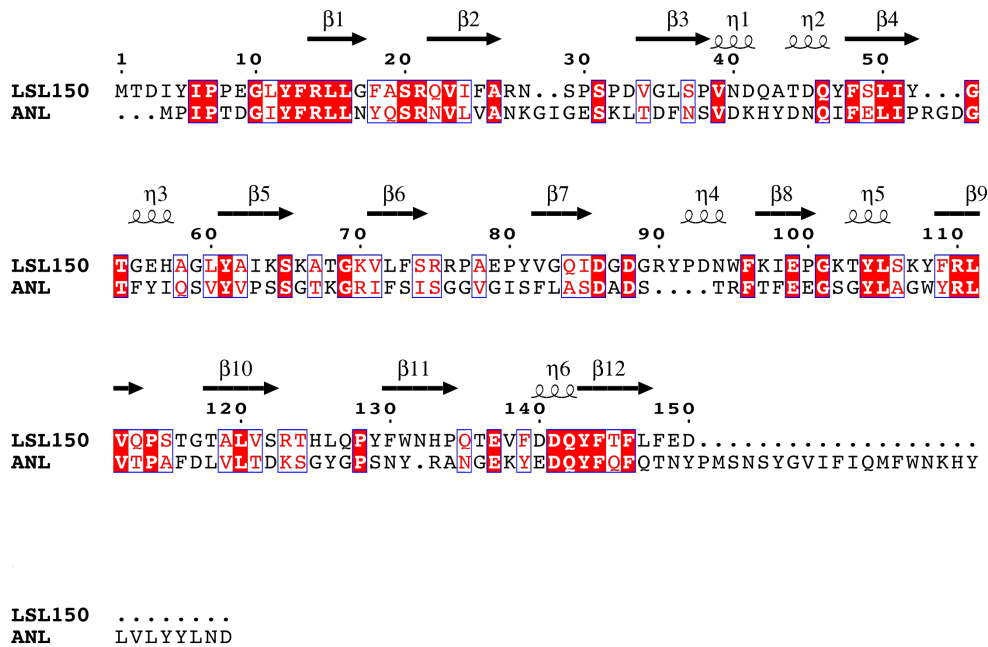
El bajo rendimiento de producción de ANL-His impidió comprobar su actividad lectina, aunque sí se pudo demostrar mediante análisis proteómico que efectivamente las bandas observadas en el *western blot* se trataban de la proteína de *A. niger*. Al igual que ocurrió con SSL, ANL-His no superó nuestro criterio de suficiencia para poder ser utilizada en alguna aplicación biotecnológica. Por ahora, debería mejorarse la expresión-purificación de la hipotética lectina de *A. niger* con el fin de aumentar el rendimiento de producción. Hasta la fecha, la actividad lectina en *A. niger* se ha estudiado mediante purificaciones parciales con precipitación con sulfato de amonio (Singh *et al.*, 2008; Singh *et al.*, 2009).

## 4.2. ANL sería un nuevo miembro de la familia de los tréboles $\beta$

A diferencia de SSL, el porcentaje de identidad de secuencia que la proteína A2R2Y9\_ASPNC comparte con LSLa es inferior al 30%, lo que dificulta incluso poder predecir su estructura tridimensional. A la vista de su secuencia (**Figura 4.45**), ANL tendría una



elevada probabilidad de adoptar un plegamiento tipo trébol  $\beta$ , ya que conserva con alguna variación el triplete clásico (QXW)<sub>3</sub> de este tipo de estructuras (Hazes, 1996): QIF (subdominio  $\alpha$ ), TRF (subdominio  $\beta$ ) y QYF (subdominio  $\gamma$ ). Un examen más minucioso de la secuencia permite ver que la proteína poseería dos sitio operativos de unión ( $\alpha$  y  $\gamma$ ) con los residuos Tyr-42 y Tyr-136 como principales responsables de la unión.

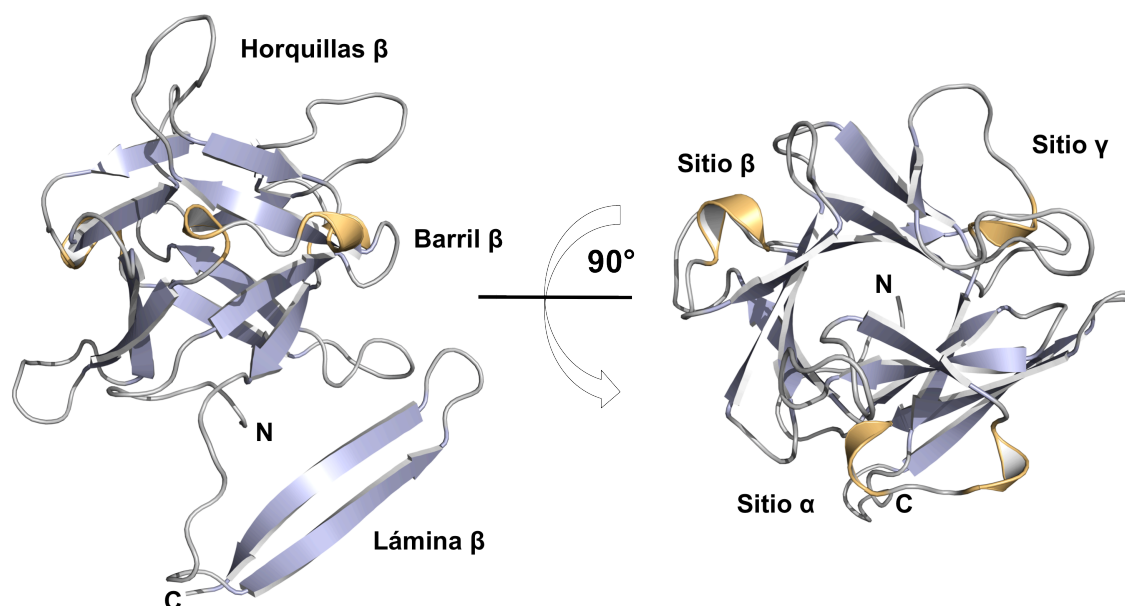


**Figura 4.45.** Alineamiento de las secuencias del módulo lectina N-terminal LSL<sub>150</sub> y la hipotética lectina ANL de *A. niger*. Los residuos idénticos en ambas secuencias se muestran sobre fondo rojo. Los residuos conservados aparecen en rojo y recuadrados en azul. Los motivos de estructura secundaria (hélices  $\alpha$  ( $\eta$ ) y hebras  $\beta$  ( $\beta$ )) de LSL<sub>150</sub> se muestran a lo largo de la secuencia.

El servidor *online* Robetta ([www.robetta.bakerlab.org](http://www.robetta.bakerlab.org)) (Kim *et al.*, 2004), que trabaja con homólogos lejanos para llevar a cabo la construcción de los modelos teóricos, permitió obtener un modelo tridimensional de la proteína ANL en base al alineamiento de secuencia con el módulo lectina LSL<sub>150</sub>. Este modelo (**Figura 4.46**) muestra que la hipotética lectina ANL adoptaría un plegamiento tipo trébol  $\beta$  muy parecido al de LSL<sub>150</sub>. Curiosamente, el extremo C-terminal, que a la vista de la secuencia, se observa que es más largo que el presente en otras lectinas de la misma familia, se organizaría formando una lámina  $\beta$  antiparalela constituida por dos hebras  $\beta$ . Este dominio adicional podría ser el responsable de los problemas de plegamiento que ocurren durante la producción recombinante de la proteína. Los diferentes modelos que arrojó el servidor Robetta se diferencian básicamente en la disposición espacial de esta región C-terminal. Esta indefinición podría sugerir que esta región tiene una elevada flexibilidad. La eliminación de este dominio extra podría suponer

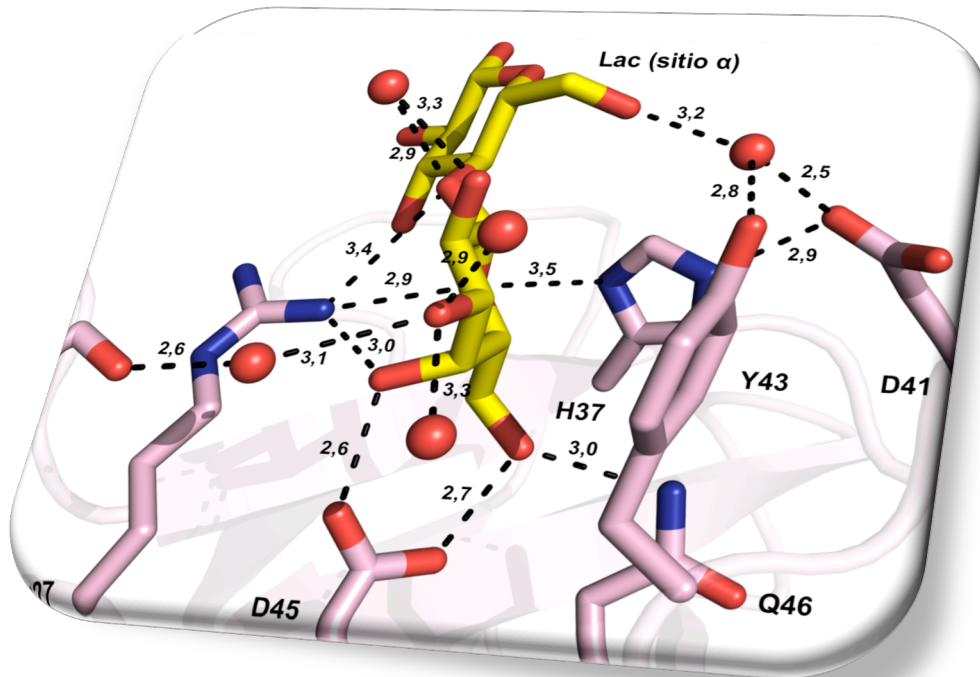


una estrategia para optimizar la producción de la lectina de forma recombinante, ayudando a aumentar la estabilidad de la misma en solución.



**Figura 4.46.** Modelo *in silico* de la estructura de ANL. Vistas frontal (izquierda) y cenital (derecha) del modelo obtenido para la estructura de ANL. En ambas estructuras las hebras  $\beta$  se representan en color *violeta* y las hélices en color *naranja*. El modelo se ha obtenido mediante el alineamiento de la secuencia de ANL con la del módulo lectina LSL<sub>150</sub> (ClustalW2, EMBL-EBI) y posterior predicción con el servidor Robetta ([www.robetta.bakerlab.org](http://www.robetta.bakerlab.org)).





## RESULTADOS Y DISCUSIÓN

Estudio de variantes  
de LSL<sub>150</sub>





## RESULTADOS Y DISCUSIÓN

### Estudio de variantes de LSL<sub>150</sub>

#### 5. Rediseño del sitio $\alpha$

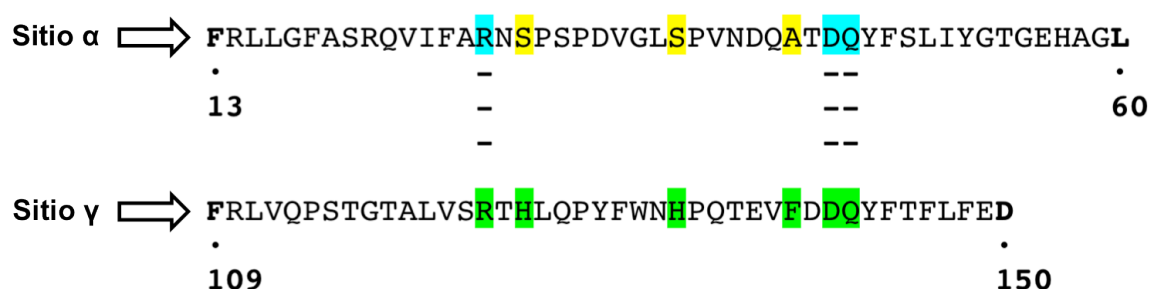
##### 5.1. Los mutantes puntuales del sitio $\alpha$ no reconstituyen la actividad lectina

###### 5.1.1. Diseño de los mutantes

Con el objetivo de esclarecer los determinantes de la unión de azúcares en la lectina de *L. sulphureus* y conseguir variantes con mayor afinidad, se llevó a cabo el diseño de una colección de mutantes puntuales de LSL<sub>150</sub>. El diseño de un sitio  $\alpha$  operativo es un modelo perfecto para realizar este estudio y validar así qué elementos contribuyen en el reconocimiento y cómo lo hacen. A la vista del estudio detallado de la unión que tiene lugar en los sitios  $\beta$  y  $\gamma$  con la molécula de lactosa, se deduce que el reconocimiento es dependiente sobre todo de dos tipos de interacciones: (i) apilamiento con aminoácidos aromáticos y (ii) enlaces de hidrógeno entre residuos polares (Arg, Asp, Asn y Gln fundamentalmente) y los grupos hidroxilo de la molécula de azúcar. Todo ello sin olvidar el papel crucial que juegan las moléculas de agua en el reconocimiento. Este comportamiento es generalizable para cualquier otro tipo de lectina, tal y como se vio en el apartado número 3 de la *Introducción*.

Si se observa en detalle la secuencia del subdominio  $\alpha$  y se compara con la del subdominio  $\gamma$  (**Figura 5.1**), donde se localiza el sitio de mayor afinidad de LSL<sub>150</sub>, se observa que este primer subdominio carece del correspondiente residuo aromático responsable del apilamiento con la cara B de la galactosa, presentando una alanina (Ala-43) en su lugar. Sin embargo, posee los residuos Arg-27, Asp-45 y Gln-46, que ocupan posiciones equivalentes a los residuos de Arg-123, Asp-141 y Gln-142, principales responsables en el sitio  $\gamma$  del establecimiento de una red de enlaces de hidrógeno con los grupos hidroxilo del anillo de galactosa del ligando. Esta comparación estructural sugeriría que la interacción hidrofóbica

con el aromático se presenta como la interacción principal que dirige la unión, promoviendo el reconocimiento del azúcar. De hecho, este escenario suele ser común en lectinas que reconocen  $\beta$ -galactósidos porque la cara B del anillo de galactosa presenta una superficie apolar más grande debido a la disposición axial del grupo OH de la posición C4 (Weis y Drickamer, 1996). Además del apilamiento con el anillo aromático, el sitio  $\gamma$  también posee dos residuos extra de histidina (His-125 e His-133) que participan en el anclaje del anillo de galactosa. Probablemente, estas histidinas, que actúan como unas pinzas, contribuirían de alguna manera a la mayor afinidad de este sitio. Estos aminoácidos tampoco aparecen en el sitio  $\alpha$ , en donde aparecen dos residuos de serina (Ser-29 y Ser-37), por lo que su incorporación podría repercutir en una mejor fijación de la lactosa y, por tanto, en la creación de un nuevo sitio de unión. Las posiciones de los residuos identificados a nivel de secuencia en el subdominio  $\alpha$  como claves para el potencial reconocimiento de una nueva molécula de ligando poseen posiciones equivalentes a nivel de la estructura tridimensional, es decir, previsiblemente la sustitución de éstos aminoácidos por un residuo equivalente a los que aparecen en los sitios  $\beta$  y/o  $\gamma$  restituiría la funcionalidad del subdominio  $\alpha$ .

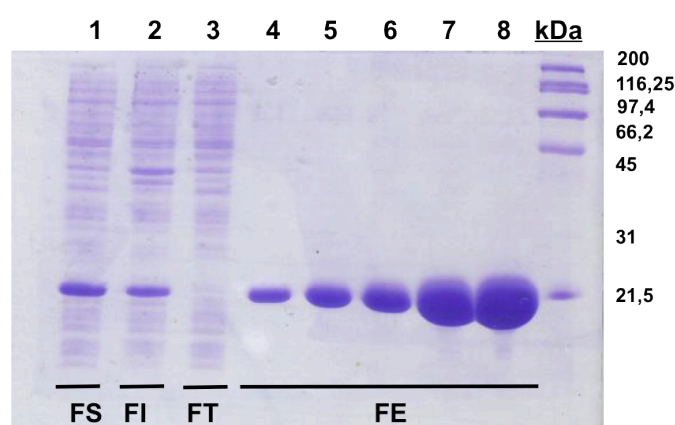


**Figura 5.1.** Alineamiento de las secuencias de aminoácidos de los subdominios  $\alpha$  y  $\gamma$  del módulo lectina N-terminal LSL<sub>150</sub>. En verde se muestran los residuos involucrados en la unión de lactosa. En azul se muestran los residuos conservados en el sitio  $\alpha$  que podrían establecer contactos con el azúcar. Por último, en amarillo se muestran las posiciones seleccionadas para el diseño de mutantes con el fin de reconstituir la capacidad de unión de azúcares en el sitio  $\alpha$ .

Por lo tanto, con objeto de convertir al sitio  $\alpha$  en un sitio operativo similar al  $\gamma$ , se diseñaron los siguientes mutantes puntuales: Ser29His (mutante  $\alpha$ 1), Ser37His (mutante  $\alpha$ 2) y Ala43Phe (mutante  $\alpha$ 3). Aunque en los tréboles  $\beta$  fúngicos el aminoácido aromático que aparece con más frecuencia en las interacciones con el ligando es una tirosina, se escogió el aminoácido de fenilalanina para hacer al sitio  $\alpha$  lo más parecido al sitio  $\gamma$  de mayor afinidad.

### 5.1.2. Purificación de los mutantes

Los tres mutantes sencillos se purificaron a homogeneidad mediante un único paso cromatográfico consistente en una cromatografía de afinidad a Sepharose® 4B, siguiendo el mismo protocolo que para LSL<sub>150</sub>. Este resultado confirma que los mutantes no pierden su capacidad para reconocer la matriz cromatográfica, razón por la cual se deduce que éstos se pliegan adecuadamente y siguen conservando su actividad lectina. En la **Figura 5.2** se muestra como ejemplo el gel de electroforesis de la purificación del mutante  $\alpha 3$ , donde se aprecia que la proteína eluye pura. Los rendimientos de producción determinados fueron de 77, 46 y 44 mg de lectina por litro de cultivo para los mutantes  $\alpha 1$ ,  $\alpha 2$  y  $\alpha 3$ , respectivamente. Al igual que ocurría con el módulo lectina silvestre, estos elevados rendimientos sugieren que los mutantes no han perdido la elevada tasa de traducción que presentaba LSL<sub>150</sub> e, igualmente, la bacteria *E. coli* sigue proporcionando el medio adecuado para su correcto plegamiento. Una vez purificadas las lectinas, ensayos de espectrometría de masas confirmaron que las mutaciones se habían introducido correctamente.



**Figura 5.2.** SDS-PAGE (15% de poliacrilamida) de la purificación del mutante  $\alpha 3$  de LSL<sub>150</sub>. En el carril 1 se muestra la fracción soluble (FS) del extracto celular, en el carril 2 la fracción insoluble (FI) y en el carril 3 la fracción no retenida a Sepharose® 4B o *flow-through* (FT). En los carriles 4-8 se muestran las fracciones del perfil de elución (FE) de la cromatografía de afinidad a Sepharose® 4B.

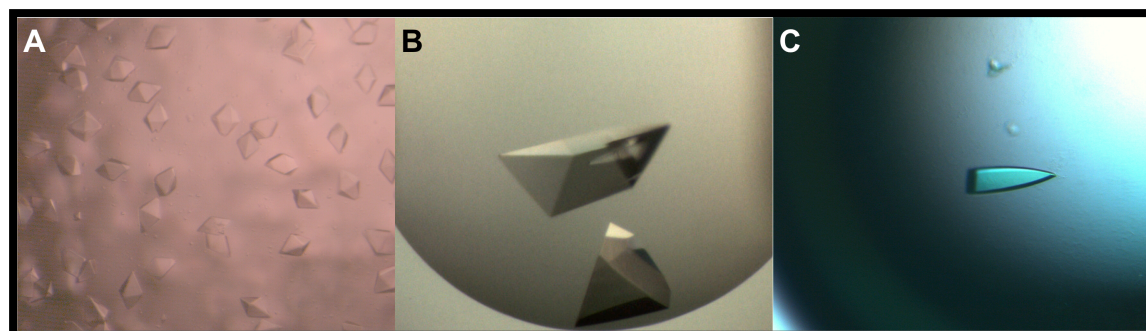
### 5.1.3. Cristalización de los mutantes

A pesar de que las lectinas sólo contienen una mutación puntual, cada una de ellas cristalizó en una condición diferente. De hecho, mutaciones puntuales que afecten a residuos de la superficie pueden modificar propiedades fisicoquímicas como la carga electrostática neta o alterar los contactos que se establecen en el cristal de la proteína silvestre, lo que conduce finalmente al uso de unas condiciones de cristalización distintas (Russo Krauss *et al.*, 2013). Sorprendentemente, el mutante  $\alpha 3$  cristaliza en el tampón en el que se encontraba disuelto (tampón TSA), lo que indica que la sustitución de la Ala-43 por una Phe aumenta considerablemente la cristalizabilidad de LSL<sub>150</sub>.

Mutante	Concentración (mg/mL)	Morfología del cristal	Condición de cristalización
$\alpha 1$	47,8	Bipirámide tetragonal	1 M sulfato de amonio y 0,1 M Tris-HCl pH 8,5
$\alpha 2$	50,3	Prisma triangular	20% (p/v) PEG 6000, 0,1 M MES pH 6,0 y 0,2 M NaCl
$\alpha 3$	17,0	Púa	25% (p/v) PEG 3350, 0,1 M Bis-tris pH 5,5 y 0,2 M NaCl

**Tabla 5.1.** Características principales de los cristales optimizados de los tres mutantes puntuales de LSL<sub>150</sub> en el sitio  $\alpha$ .

Los cristales de los tres mutantes (**Figura 5.3**) se obtuvieron mediante cocrystalización con 0,2 M lactosa con la intención de comprobar a partir de la estructura si el sitio  $\alpha$  es capaz de unir carbohidratos. Inicialmente, en un plazo de una a dos semanas, se obtuvieron multitud de cristales en diversas condiciones con los *screening* comerciales, preparados en microplaca con ayuda del robot de cristalización. Estos cristales se optimizaron, posteriormente, usando la técnica de cristalización por difusión de vapor en gota colgante a una temperatura de 18 °C (**Tabla 5.1**).



**Figura 5.3.** Cristales optimizados de los mutantes sencillos del sitio  $\alpha$  de LSL<sub>150</sub> obtenidos mediante difusión de vapor en gota colgante en presencia de 0,2 M lactosa a 18 °C. **A.** Cristales del mutante  $\alpha 1$  (Ser29His) aparecidos en 0,1 M Tris-HCl pH 8,5 y 1 M sulfato de amonio. **B.** Cristales del mutante  $\alpha 2$  (Ser37His) aparecidos en 20% (p/v) PEG 6000, 0,1 M MES pH 6,0 y 0,2 M NaCl. **C.** Fotografía bajo luz polarizada del cristal del mutante  $\alpha 3$  (Ala43Phe) aparecido en 25% (p/v) PEG 3350, 0,1 M Bis-tris pH 5,5 y 0,2 M NaCl.

#### 5.1.4. Resolución estructural de los mutantes

Los cristales de cada uno de los mutantes fueron medidos en el ESRF (Grenoble, Francia). El set de datos correspondiente al mutante  $\alpha 1$  fue recogido en la línea ID23-1 a una longitud de onda de 1,067 Å, mientras que los datos de difracción de los otros dos mutantes fueron tomados en la línea ID23-2 a una longitud de onda de 0,873 Å. En los tres casos, los cristales se crioprotegieron con un 25% (v/v) glicerol y se midieron a 100 K, recogiendo 200 imágenes con un ángulo de rotación de 1°.



Los datos de difracción (**Tabla 5.2**) se procesaron en los tres casos con los programas iMosflm (Leslie, 2006) y SCALA (Evans, 2006), incluidos en el paquete de software CCP4 (Cowtan *et al.*, 2011; Winn *et al.*, 2011). Los cristales demostraron ser de muy buena calidad, dado que los patrones de difracción mostraban reflexiones a alta resolución (mayores de 2 Å). La resolución estructural se llevó a cabo mediante reemplazo molecular, utilizando el programa MOLREP (Vagin y Teplyakov, 2010) y usando como modelo de reemplazo la estructura del módulo lectina LSL<sub>150</sub>. Los cristales de los mutantes  $\alpha 2$  y  $\alpha 3$  conservaban el mismo grupo espacial que el obtenido para el cristal del complejo de LSL<sub>150</sub> con lactosa (un grupo espacial  $P3_2$  con una molécula en la unidad asimétrica). En cambio, los cristales del mutante  $\alpha 1$  eran completamente diferentes, pues eran tetragonales (grupo espacial  $P4_12_12_1$ ) y con dos moléculas de proteína en la unidad asimétrica.

Parámetros	LSL <sub>150</sub> - $\alpha 1$ :lactosa	LSL <sub>150</sub> - $\alpha 2$ :lactosa	LSL <sub>150</sub> - $\alpha 3$ :lactosa
Recogida de datos y procesamiento			
Grupo espacial	$P4_12_12_1$	$P3_2$	$P3_2$
Moléculas en la UA	2	1	1
Celda unidad (Å)	106,61, 106,61, 88,90	62,72, 62,72, 36,27	62,39, 62,39, 36,68
Longitud de onda (Å)	1,067	0,873	0,873
Resolución (Å)	68–2,10	22–1,48	28–1,34
Reflexiones totales	363.227	167.928	228.778
Reflexiones únicas	30.182	26.594	34.838
Compleitud	99,6 (99,6)	100 (100)	100 (100)
Multiplicidad	12 (11,7)	6,3 (6,1)	6,4 (6,2)
$R_{merge}$ (%)	8,1 (41,8)	7,1 (48,2)	5,5 (47,8)
$I/\sigma(I)$	16,9 (5,5)	14,4 (3,5)	17,8 (3,7)
Refinamiento			
$R_{work}/R_{free}$	0,171/0,214	0,163/0,197	0,160/0,184
RMSD de los enlaces (Å)	0,008	0,007	0,006
RMSD de los ángulos (°)	1,250	1,295	1,230
Factor $B$ medio (Å <sup>2</sup> )	48,72	18,86	20,08
Número de átomos			
Proteína	2.440	1.215	1.244
Glicerol	0	18	6
Lactosa	46	69	69
Agua	154	234	250
Mapa de Ramachandran			
Favorecidos (%)	95,2	98,0	98,0
Permitidos (%)	4,1	1,4	2,0
Prohibidos (%)	0,7	0,6	0,0

\*Los valores de la concha de mayor resolución se muestran en paréntesis.

**Tabla 5.2.** Datos de difracción y estadísticas del refinamiento de las estructuras de los mutantes  $\alpha 1$  (Ser29His),  $\alpha 2$  (Ser37His) y  $\alpha 3$  (Ala43Phe) de LSL<sub>150</sub> en complejo con lactosa.

El refinamiento de los modelos se llevó a cabo con PHENIX (Adams *et al.*, 2010). Los mutantes conservan prácticamente la misma estructura que la proteína silvestre, por lo que la reconstrucción se ciñó a la zona donde se habían realizado las mutaciones. Después de aplicar refinamiento anisótropo y TLS, los índices de validación  $R_{work}/R_{free}$  alcanzaron valores finales de 0,171/0,214, 0,163/0,197 y 0,160/0,184 para los mutantes  $\alpha 1$ ,  $\alpha 2$  y  $\alpha 3$ , respectivamente. El resto de datos estadísticos acerca del refinamiento quedan recogidos en la **Tabla 5.2**.

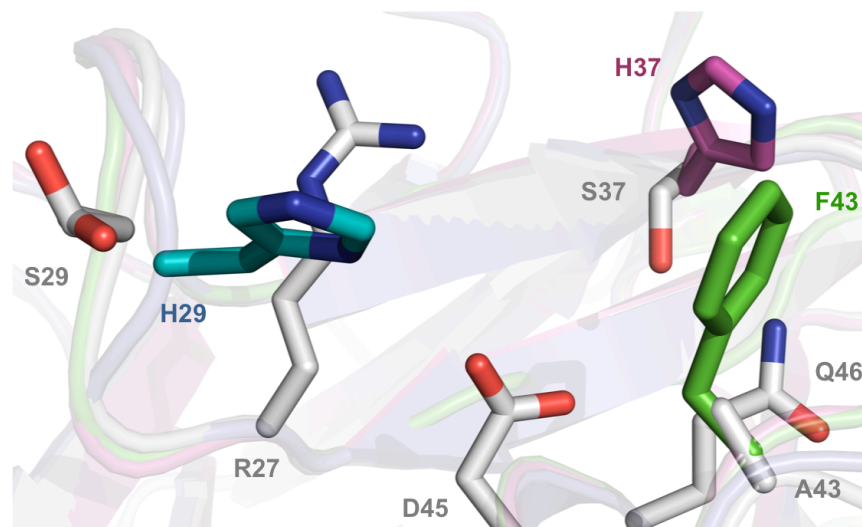
### 5.1.5. Análisis estructural del sitio $\alpha$

El análisis de la estructura de los mutantes en comparación con la proteína silvestre demuestra que la estructura terciaria se mantiene a pesar de las mutaciones introducidas. De hecho, el RMSD de los C $\alpha$  entre los mutantes y LSL<sub>150</sub> es inferior a 1 Å (0,53, 0,42 y 0,38 Å para los mutantes  $\alpha 1$ ,  $\alpha 2$  y  $\alpha 3$ , respectivamente).

Con respecto a los sitios de unión de azúcares  $\beta$  y  $\gamma$  de los mutantes: el sitio  $\gamma$  sigue siendo funcional en las tres variantes. Las interacciones que se establecen en este sitio son equivalentes a las descritas anteriormente en la sección 1.2.3 del capítulo IV de *Resultados y Discusión* para la proteína silvestre. Por otro lado, el sitio  $\beta$  no aparece ocupado en el mutante  $\alpha 1$ , mientras que en los otros dos mutantes se reproduce el mismo escenario que en LSL<sub>150</sub>. De hecho, en una de las dos moléculas que constituyen la unidad asimétrica del cristal en el mutante  $\alpha 1$  se observa que el extremo N-terminal de una molécula relacionada por simetría se inserta en el sitio  $\beta$ , lo que impide, por tanto, la unión de la molécula de ligando. En la otra molécula de la unidad asimétrica, en cambio, sí que hay espacio para que se produzca la unión. Probablemente, este resultado esté relacionado estrechamente con la condición de cristalización. Como se había indicado previamente, el sitio  $\beta$  en LSL<sub>150</sub> es un sitio de muy baja afinidad y, por tanto, podría ser más susceptible a aparecer desocupado bajo determinadas condiciones. En este sentido, la estructura de LSLa (Mancheño *et al.*, 2005) también fue resuelta en presencia de una única molécula de lactosa en el sitio  $\gamma$ , lo que es consistente con la menor capacidad de unión del sitio  $\beta$ .

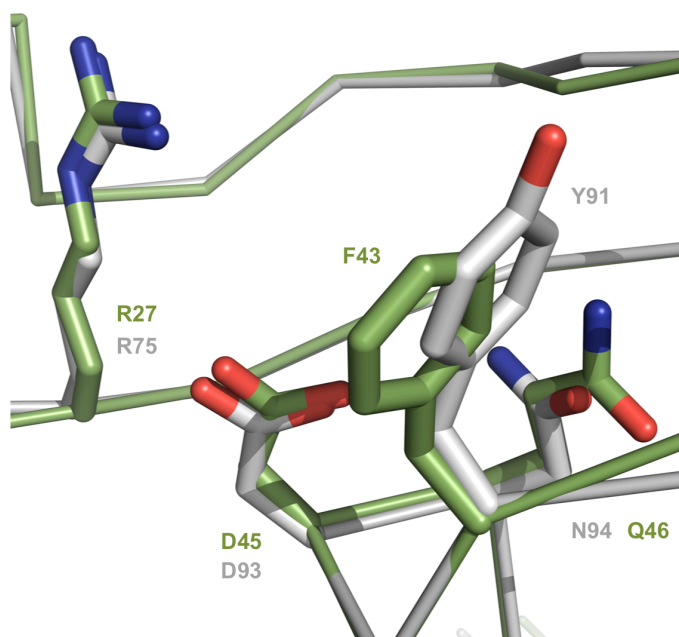
Ninguna de las mutaciones puntuales que se efectuaron sobre el sitio  $\alpha$  condujeron a hacer de él un sitio operativo. La incorporación individual de los nuevos aminoácidos, por tanto, no es condición suficiente para que la lactosa se fije en este entorno. La superposición del sitio  $\alpha$  de los tres mutantes (**Figura 5.4**) demuestra que aparentemente los residuos introducidos, en su conjunto, conformarían un sitio estructuralmente similar al sitio  $\gamma$ ,

por lo que se hace necesario evaluar cuál sería la contribución conjunta de los tres aminoácidos (apartado 5.2 del capítulo V de *Resultados y Discusión*). En el mutante  $\alpha 1$  sí se identificó, en cambio, una molécula de glicerol en el entorno del sitio  $\alpha$ , cuya procedencia probablemente sea la solución crioprotectora.



**Figura 5.4.** Superposición en el entorno del sitio  $\alpha$  de las estructuras de LSL<sub>150</sub> y de sus correspondientes mutantes puntuales  $\alpha 1$ ,  $\alpha 2$  y  $\alpha 3$ . La posición de los residuos de LSL<sub>150</sub> (gris) se compara con la de los residuos introducidos mediante mutagénesis dirigida: His29 (azul) del mutante  $\alpha 1$ , H37 (fucsia) del mutante  $\alpha 2$  y F43 (verde) del mutante  $\alpha 3$ .

El sitio  $\alpha$  del mutante  $\alpha 3$  es estructuralmente similar al sitio  $\beta$  de LSL<sub>150</sub>. La comparación de ambos sitios (**Figura 5.5**) revela que los aminoácidos Arg-27, Phe-43, Asp-45 y Gln-46 ocupan las mismas posiciones que los residuos Arg-75, Tyr-91, Asp-93 y Asn-94, implicados directamente en el reconocimiento del disacárido. Se podría considerar que la diferencia en el tipo de residuo aromático es esencial para la unión de lactosa. Sin embargo, la presencia de un residuo de Phe en el sitio  $\gamma$  no evita que se lleve a cabo la unión. De hecho, aunque la Tyr-91 efectivamente está implicada en la interacción con lactosa a través de una molécula de agua y el grupo hidroxilo del anillo (apartado 1.2.3 del capítulo IV de *Resultados y Discusión*), tal y como se verá en la próxima sección, el grupo hidroxilo no es *per se* esencial para la unión de lactosa. Posiblemente es más relevante la contribución de las moléculas de agua en el reconocimiento de los azúcares (Weis y Drickamer, 1996; Tschampel y Woods, 2003). Por esta razón, muchos intentos de (re)diseñar/optimar sitios de unión acaban siendo intentos frustrados principalmente por la falta de control que se posee sobre el papel del solvente en las interacciones con los ligandos (Toone, 1994).



**Figura 5.5.** Superposición de los sitios  $\alpha$  (verde) y  $\beta$  (gris) del mutante  $\alpha 3$  de LSL<sub>150</sub>. Se representan los residuos claves en la unión de la lactosa en el sitio  $\beta$  y sus homólogos en el sitio  $\alpha$ .

## 5.2. Los mutantes triples hacen funcional al subdominio $\alpha$

### 5.2.1. Diseño, purificación, cristalización y resolución estructural de los mutantes triples

Teniendo en cuenta el análisis de la secuencia de aminoácidos de los tres subdominios de LSL<sub>150</sub> y los resultados estructurales obtenidos con los mutantes puntuales del sitio  $\alpha$ , se llevó a cabo el diseño de dos nuevos mutantes del módulo lectina. Estos mutantes incorporan al mismo tiempo las tres mutaciones que se habían estudiado previamente. Se diseñaron dos variantes: una con la mutación Ala43Phe y otra con el cambio Ala43Tyr, por lo que se planificaron los mutantes triples  $\alpha F$  (Ser29His, Ser37His, Ala43Phe) y  $\alpha Y$  (Ser29His, Ser37His, Ala43Tyr). La incorporación de los aminoácidos aromáticos distintos permitirá evaluar cómo afecta la naturaleza de cada uno de éstos a la unión. Por un lado, Tyr está presente en el sitio  $\beta$  de LSL<sub>150</sub>, además de ser el que con más frecuencia aparece en los sitios de unión de las lectinas fúngicas con estructura tipo trébol  $\beta$  (Tabla 4.3), mientras que Phe lo está en el sitio  $\gamma$ .

La expresión y purificación de los dos triples mutantes se llevó a cabo con éxito siguiendo el mismo protocolo de la purificación del módulo lectina LSL<sub>150</sub> y sus respectivos mutantes sencillos basados en la cromatografía de afinidad a Sepharose® 4B. Después de este paso cromatográfico, las lectinas se obtuvieron puras con unos rendimientos de ~40 mg por litro de cultivo. Aunque se siguen observando altos rendimientos de produc-

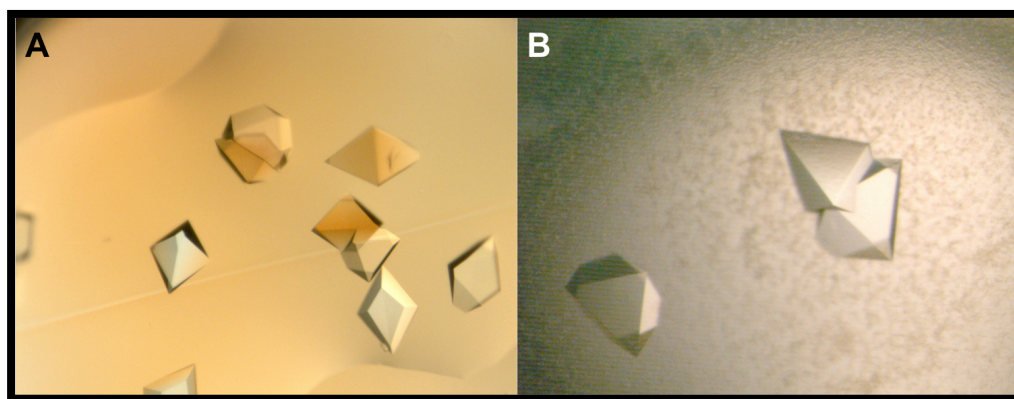
ción, estos son aproximadamente la mitad del registrado en la purificación de la proteína silvestre.

Parámetros	LSL <sub>150</sub> - $\alpha$ F:lactosa	LSL <sub>150</sub> - $\alpha$ Y:lactosa
Recogida de datos y procesamiento		
Grupo espacial	$P4_12_12$	$P2_1$
Moléculas en la UA	1	4
Celda unidad (Å)	69,38, 69,38, 67,46	66,46, 70,09, 69,98
Longitud de onda (Å)	0,939	0,979
Resolución (Å)	35–1,10	48–1,55
Reflexiones totales	982.572	627.181
Reflexiones únicas	66.758	91.778
Complejidad	99,3 (93,1)	98,2 (97,0)
Multiplicidad	14,7 (9,4)	6,8 (6,5)
$R_{merge}$ (%)	7,9 (60,1)	7,5 (71,7)
$I/\sigma(I)$	19,0 (3,5)	13,2 (2,8)
Refinamiento		
$R_{work}/R_{free}$	0,146/0,162	0,163/0,192
RMSD de los enlaces (Å)	0,012	0,013
RMSD de los ángulos (°)	1,410	1,240
Factor $B$ medio (Å <sup>2</sup> )	14,40	20,60
Número de átomos		
Proteína	1.221	4.832
Lactosa	46	276
GGG	34	---
Agua	254	715
Mapa de Ramachandran		
Favorecidos (%)	99,0	98,3
Permitidos (%)	1,0	1,7
Prohibidos (%)	0	0

\*Los valores de la concha de mayor resolución se muestran en paréntesis.

**Tabla 5.3.** Datos de difracción y estadísticas del refinamiento de las estructuras de los mutantes  $\alpha$ F (Ser29His, Ser72His, Ala43Phe) y  $\alpha$ Y (Ser29His, Ser72His, Ala43Tyr) de LSL<sub>150</sub> en complejo con lactosa.

Una vez los mutantes fueron purificados, se realizaron experimentos de cristalización. Al igual que ocurrió con los mutantes puntuales, los dos mutantes triples cristalizaron en condiciones diferentes a las identificadas para los otros casos, lo que vuelve a indicar que cambios en la superficie afectan considerablemente a la cristalizabilidad. Ambos mutantes se cristalizaron en presencia de 0,2 M lactosa con el fin de evaluar la funcionalidad de los nuevos sitios  $\alpha$ . Por un lado, el mutante  $\alpha$ F cristalizó a ~60 mg/mL en 3,5 M formiato de sodio pH 7,0. Los cristales aparecieron en microplaca con la disolución comercial y al cabo de un mes de tiempo (**Figura 5.6A**). Por otro lado, los mejores cristales del mutante  $\alpha$ Y se obtuvieron a una concentración de 22 mg/mL en la condición 25% ( $p/v$ ) PEG 3350, 0,1 M HEPES pH 7,5 y 0,2 M MgCl<sub>2</sub>, optimizada en el laboratorio (**Figura 5.6B**).



**Figura 5.6.** Cristales optimizados de los mutantes triples del sitio  $\alpha$  de LSL<sub>150</sub> obtenidos mediante difusión de vapor en gota colgante en presencia de 0,2 M lactosa a 18 °C. **A.** Fotografía de los cristales del mutante  $\alpha$ F (Ser29His, Ser37His, Ala43Phe) bajo luz polarizada aparecidos en 3,5 M formiato de sodio pH 7,0. **B.** Cristales del mutante  $\alpha$ Y (Ser29His, Ser37His, Ala43Tyr) aparecidos en 25% (p/v) PEG 3350, 0,1 M HEPES pH 7,5 y 0,2 M MgCl<sub>2</sub>.

Los cristales del mutante  $\alpha$ F se midieron en la línea ID14-1 del ESRF (Grenoble, Francia) a una longitud de onda de 0,939 Å y usando un detector ADSC Q315R. Los cristales del mutante  $\alpha$ Y, sin embargo, fueron medidos en la línea BL13 (XALOC) del sincrotrón ALBA (Sardañola del Vallés, España) con una longitud de onda de 0,979 Å y usando un detector PILATUS 6M. En los dos casos, la recogida de datos se hizo a 100 K y con un ángulo de giro por imagen de 1°. Los cristales de ambos mutantes se crioprotegieron con 20% (v/v) glicerol antes de efectuarse las medidas de difracción de rayos X.

Aunque los cristales de ambos mutantes son morfológicamente parecidos, el análisis de los datos cristalográficos reveló que los grupos espaciales a los que pertenece cada uno son diferentes. Los cristales del mutante  $\alpha$ F pertenecen al grupo espacial tetragonal  $P4_12_12$  y tienen una única molécula en la unidad asimétrica, mientras que el mutante  $\alpha$ Y presenta un grupo espacial  $P2_1$  y posee cuatro moléculas de proteína en la unidad asimétrica. Los datos de difracción obtenidos fueron de muy buena calidad y elevada resolución, lo que permitió que las estructuras de los mutantes  $\alpha$ F y  $\alpha$ Y se resolvieran a una resolución de 1,10 y 1,55 Å, respectivamente. Al igual que en el caso de los mutantes puntuales, las estructuras de los triples mutantes se resolvieron mediante reemplazo molecular, usando la estructura de LSL<sub>150</sub> como modelo. El refinamiento, llevado a cabo con PHENIX (Adams *et al.*, 2010), fue diferente en ambos casos. La mayor resolución del mutante  $\alpha$ F permitió llevar a cabo refinamiento anisótropo e introducir átomos de hidrógeno en el modelo final (sólo en la proteínas y no en los ligandos ni en las aguas). Por otro lado, en el refinamiento del mutante  $\alpha$ Y se aplicó refinamiento TLS. Finalmente, las dos estructuras se obtuvieron

con unos valores de los factores  $R_{work}/R_{free}$  de 0,146/0,162 y 0,163/0,192 para los mutantes  $\alpha F$  y  $\alpha Y$ , respectivamente. El resto de parámetros cristalográficos obtenidos durante el procesamiento se recogen en la **Tabla 5.3**.

### 5.2.2. Análisis estructural de la unión de lactosa en el nuevo sitio $\alpha$

A la vista de los modelos finales de las estructuras de los mutantes triples, se observa que la incorporación de las tres mutaciones no afecta a la estructura tridimensional final, que se mantiene prácticamente intacta, es decir, igual a la de la proteína silvestre (el valor medio del RMSD de los  $C\alpha$  considerando todos los átomos de la proteína es de 0,24 tanto para el mutante  $\alpha F$  como para el  $\alpha Y$ ).

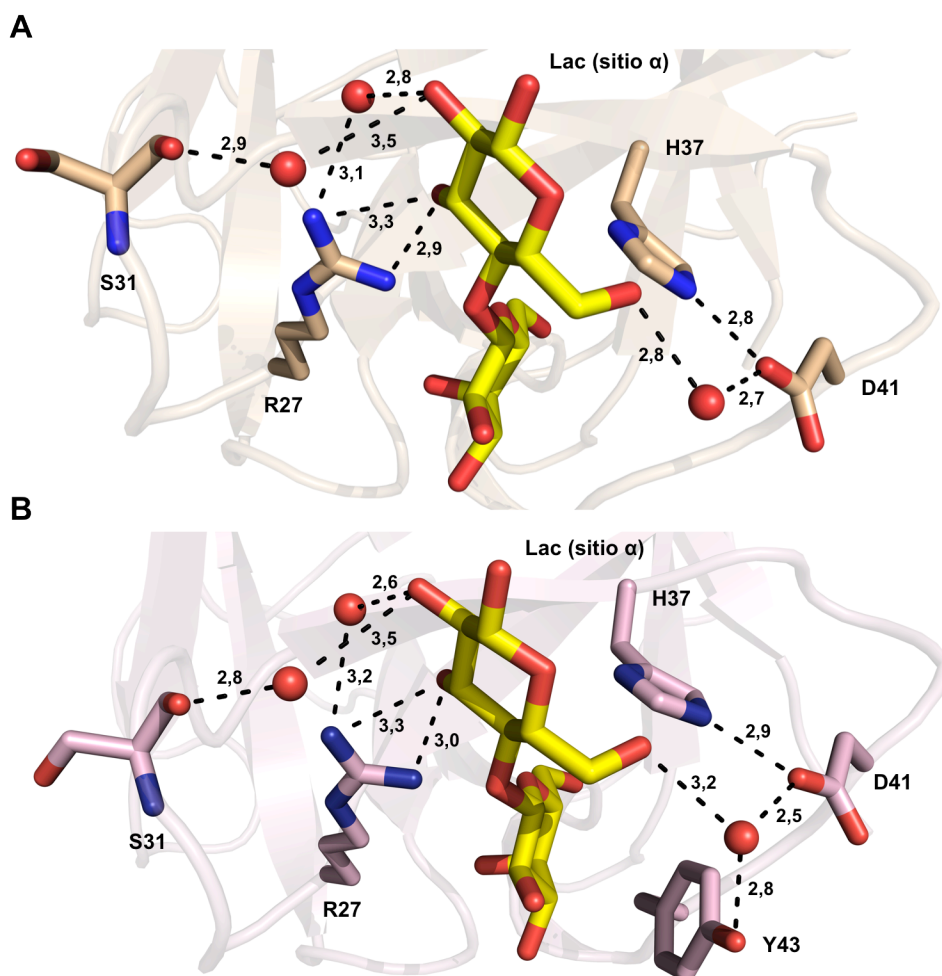
El análisis cristalográfico de los sitios de unión reveló que los sitios  $\beta$  y  $\gamma$  conservan la capacidad de unir carbohidratos en ambos mutantes, así como que el sitio  $\alpha$  rediseñado es capaz de unir lactosa, independientemente de si el residuo aromático es una Phe o una Tyr. Cabe destacar que a diferencia del sitio de unión  $\gamma$ , donde las densidades electrónicas de los ligandos están perfectamente definidas, en los sitios  $\alpha$  y  $\beta$ , la densidad electrónica es de peor calidad en torno al ligando, sobre todo en la zona del residuo de glucosa. Al igual que ocurre en la proteína silvestre, el residuo de glucosa está más expuesto a la superficie, presenta menos contactos con la lectina y, por tanto, posee mayor flexibilidad (factores  $B$  más altos). Este hecho, unido a una menor ocupación, que quizás esté dando una idea de su menor afinidad, podrían justificar la peor calidad de los mapas de densidad electrónica en estas regiones.

A diferencia de lo que se observa en el mapa de densidad electrónica de LSL<sub>150</sub>, los mapas de ambos mutantes triples no permitieron analizar la especificidad por un anómero u otro en torno al sitio  $\alpha$ . En este caso, usamos el anómero  $\beta$  para el estudio detallado de las interacciones proteína-azúcar, puesto que es el anómero que aparece en la naturaleza en mayor proporción y el que mayor ocupación presenta tanto en las estructuras de LSL<sub>150</sub> como de LBL<sub>152</sub>.

Las interacciones que se establecen en los sitios de unión  $\alpha$  de los dos mutantes triples son prácticamente idénticas entre sí con excepción de las interacciones que involucran a los residuos aromáticos. Con respecto a las interacciones que se forman con el anillo de glucosa (**Figura 5.7**), en primer lugar se observa el importante papel que juega el residuo de Arg-27, que forma dos enlaces de hidrógeno entre los nitrógenos  $N\eta$  del grupo guanidinio y el oxígeno O3' del ligando. Este aminoácido también reconoce a una molécula de agua



que a su vez forma un puente con el O2'. El oxígeno O2' además reconoce a otra molécula de agua que forma un enlace de hidrógeno al mismo tiempo con el oxígeno carbonílico de la cadena principal (residuo Ser-31). Por otro lado, el oxígeno O6' participa en una interacción con una molécula de agua que a su vez es reconocida por el oxígeno Oδ del Asp-41. En el caso del mutante  $\alpha$ Y, esta molécula de agua es reconocida por un segundo aminoácido, la Tyr-43, que tiene la capacidad de formar un enlace de hidrógeno a través del grupo hidroxilo del anillo aromático.



**Figura 5.7.** Vista detallada de las interacciones específicas entre el anillo de glucosa del anómero  $\beta$  de la lactosa y los residuos del sitio de unión  $\alpha$  de los mutantes LSL<sub>150</sub>- $\alpha$ F (A) y LSL<sub>150</sub>- $\alpha$ Y (B). Las moléculas de agua se representan mediante *esferas* de color *rojo*. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea discontinua* y las medidas de las distancias vienen dadas en *ångströms* (Å).

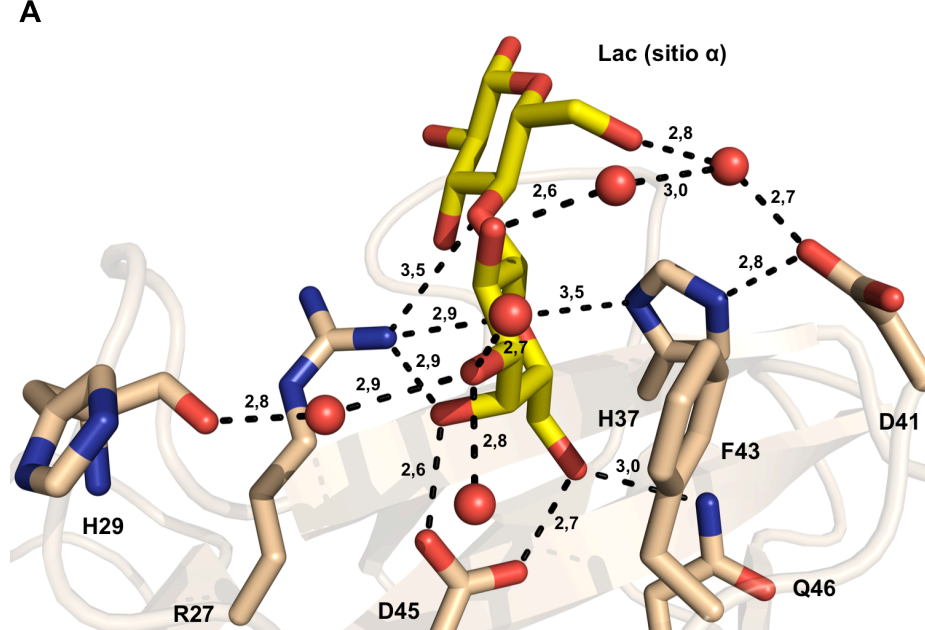
Las interacciones que se observan entre el residuo de glucosa y el sitio de unión no son nuevas, es decir, ya habían sido descritas anteriormente para los otros sitios de unión caracterizados hasta el momento. Por ejemplo, la molécula de agua que actúa como puente entre el residuo de Asp-41 y el oxígeno O6' aparece también en el sitio  $\gamma$  de la lectina silvestre entre el Glu-137 y el oxígeno O6' de la lactosa (Figura 4.7B). Además, en este último



caso, la His-133 forma un enlace de hidrógeno con el Glu-137 que ayuda a mantener una correcta orientación de la cadena lateral del residuo ácido. Esta interacción auxiliar aparece también en el nuevo sitio  $\alpha$  gracias a la introducción del residuo de His-37. La His-37, por tanto, forma un enlace de hidrógeno a través del nitrógeno N $\epsilon$ -2 con uno de los oxígenos de la cadena lateral del Asp-41. Este papel estabilizador de la His-37 podría ser clave para la unión y, por tanto, justificaría su presencia en el sitio de unión de azúcares. Los enlaces que se observan en el aminoácido Arg-27, por otro lado, tampoco son novedosos, puesto que la presencia de un residuo de Arg es una constante en los sitios de unión tanto de LSL<sub>150</sub> como de LBL<sub>152</sub>. Finalmente, la participación de la Ser-31 en la formación de un puente con el O2' a través de una molécula de agua ya había sido descrita previamente en el sitio de unión  $\alpha$  de la lectina homóloga LBL<sub>152</sub>.

Mayor es el número de contactos que se ven entre el anillo de galactosa y el sitio  $\alpha$  (**Figura 5.8**). En este caso, aunque los enlaces de hidrógeno que se observan en ambos mutantes son prácticamente idénticos, la diferencia principal se debe al apilamiento que se establece con el residuo aromático. Sin embargo, ambos ocupan la misma posición, disponiéndose a una distancia de 3,9 Å de la cara B del anillo de galactosa. Este resultado indica que, a pesar de las diferencias entre las cadenas laterales de la fenilalanina y la tirosina, no existen diferencias apreciables que afecten al reconocimiento de la lactosa en los nuevos sitios de unión  $\alpha$ . En relación con el resto de interacciones, la Arg-27 vuelve a participar activamente en la unión, estableciendo enlaces de hidrógeno con el oxígeno del enlace glucosídico, el oxígeno del anillo y el oxígeno O4. El O4 y el O6, a su vez, forman dos enlaces de hidrógeno con el Asp-45, que se coloca justamente debajo del anillo de galactosa, al fondo de la cavidad. El O6 también tiene la capacidad de reconocer al nitrógeno N $\epsilon$ -2 de la Gln-46. Paralelamente, el oxígeno del anillo forma un enlace de hidrógeno con la His-37, lo que indica que este residuo es importante en el reconocimiento. Por último, cabe destacar que los oxígenos O2 y O3 no reconocen directamente ningún aminoácido, pero establecen diferentes enlaces con moléculas de agua de su entorno. De esta manera, el O3 forma un puente con el oxígeno peptídico de la His-29 a través de una molécula de agua. En el caso particular del mutante  $\alpha$ F, el O2 forma un enlace de hidrógeno con una molécula de agua que suple el papel del grupo hidroxilo de la Tyr-43 del mutante  $\alpha$ Y porque está reconociendo a la molécula de agua localizada entre el oxígeno O6' y el Asp-41.

Al igual que ocurre con las interacciones identificadas en el anillo de glucosa, todas las interacciones observadas que involucran al anillo de galactosa ya habían sido descritas



**Figura 5.8.** Vista detallada de las interacciones específicas entre el anillo de galactosa del anómero  $\beta$  de la lactosa y los residuos del sitio de unión  $\alpha$  de los mutantes LSL<sub>150</sub>- $\alpha$ F (**A**) y LSL<sub>150</sub>- $\alpha$ Y (**B**). Las moléculas de agua se representan mediante *esferas* de color rojo. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en *línea discontinua* y las medidas de las distancias vienen dadas en *ångströms* (Å).

La comparación de las interacciones que se establecen en los tres sitios de unión ( $\alpha$ ,  $\beta$  y  $\gamma$ ) arroja algunos resultados interesantes (**Tabla 5.4**). Con respecto al número de interacciones que se establecen, el sitio  $\gamma$  cuenta con más interacciones que el sitio  $\alpha$ , y éste a su vez con más que el sitio  $\beta$ . Fundamentalmente, la principal diferencia recae en la disposición de las pinzas de histidinas (formadas por las His-125 y 133) que fijan el anillo de galactosa en el sitio  $\gamma$ . Esto probablemente se traduciría en una interacción más fuerte con el ligando en este sitio de unión. Estas pinzas no existen en el sitio  $\beta$  y no se llegan a formar en el sitio  $\alpha$  de ninguno de los dos mutantes triples de LSL<sub>150</sub> por la orientación inesperada que adopta la His-29. La presencia de residuos de His en los sitios de unión no es tan frecuente en los tréboles  $\beta$ . Por ejemplo, entre las lectinas fúngicas con este tipo de plegamiento, solamente, RSA posee una His (His-40) en el sitio  $\alpha$  que es capaz de formar un enlace de hidrógeno con el oxígeno O3 de la galactosa (Skamnaki *et al.*, 2013).

	LSL <sub>150</sub> - $\alpha$ F/ $\alpha$ Y		
	$\alpha$	$\beta$	$\gamma$
<b>Especificidad</b>	$\beta$ -galactósidos	$\beta$ -galactósidos	$\beta$ -galactósidos
<b>Aromático</b>	Phe43/Tyr43	Tyr91	Tyr142
<b>Ligando-O2</b>	—	—	—
<b>Ligando-O3</b>	—	—	—
<b>Ligando-O4</b>	Arg27 y Asp45	Arg75 y Asp93	Arg123, His125 y Asp141
<b>Ligando-O6</b>	Asp45 y Gln46	Asp93 y Asn94	Asp141 y Gln142
<b>Ligando-O<sub>anillo</sub></b>	His37	—	His133

**Tabla 5.4.** Comparación de las interacciones que se establecen en los sitios operativos de unión de azúcar de los mutantes triples LSL<sub>150</sub>- $\alpha$ F (Ser29His, Ser72His, Ala43Phe) y LSL<sub>150</sub>- $\alpha$ Y (Ser29His, Ser72His, Ala43Tyr). Solo se han analizado los residuos que interaccionan directamente con el anillo de galactosa.

Es interesante destacar también el papel del resto de aminoácidos que conforman los sitios de unión de azúcares. Mientras que en el sitio  $\gamma$  todos los aminoácidos están participando en la unión, ya sea reconociendo al anillo de glucosa o al de galactosa directamente o a través de moléculas de agua; en los sitios  $\alpha$  y  $\beta$  hay residuos de carácter apolar cuyo papel en la unión de los ligandos está aún por esclarecer. En el caso del sitio  $\beta$ , se observan las cadenas voluminosas de la Phe73 y la Ile85; y en el caso del sitio  $\alpha$ , la Phe25. Aunque estos aminoácidos no impiden estéricamente la unión, podrían dar lugar a la aparición de repulsiones debido a la naturaleza polar de los azúcares. En este sentido, la repulsión sería menor en el caso del sitio  $\alpha$ . Aunque, desde otro punto de vista, podrían intensificar la

unión debido a que la presencia de lactosa evita que los residuos de carácter apolar queden expuestos al solvente, lo cual está favorecido termodinámicamente.

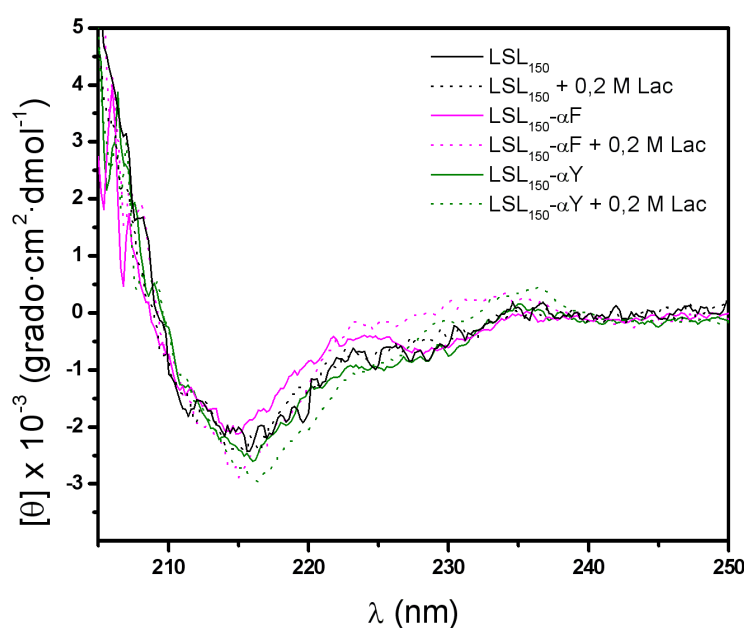
En conclusión, los resultados aquí mostrados advierten de la complejidad que subyace a la interacción entre un azúcar relativamente simple, como la lactosa, y una lectina. Aunque la unión pueda estar dirigida por la interacción que se establece entre los protones alifáticos de la cara B de la galactosa y la nube  $\pi$  de electrones del residuo aromático (Weis y Drickamer, 1996), el establecimiento de una red de moléculas de agua, así como de un número suficiente de contactos para fijar ambas unidades del disacárido, son críticos para que un sitio potencial de unión se vuelva operativo. De hecho, un experimento similar fue llevado a cabo sobre la lectina CNL con la intención de reconstruir un sitio de unión en el subdominio  $\gamma$  que calcase las características del sitio operativo  $\alpha$ . En este caso, el mutante triple Thr116Asp, Asp120Ser y Asp132Asn, diseñado para tal fin, es incapaz de unir ningún ligando en el nuevo sitio. Aparentemente, la estructura del sitio  $\gamma$  muestra cómo los aminoácidos no adquieren la misma orientación que en el sitio  $\alpha$ , por lo que no pueden formar correctamente los respectivos enlaces de hidrógeno (Pohleven *et al.*, 2012). Este resultado es consistente con la necesidad de que exista un microentorno que favorezca la disposición correcta de las cadenas laterales y asegure la permanencia de las moléculas de agua clave para el reconocimiento.

Creemos que otro aspecto interesante a la vista de nuestros resultados, reveladores de una complejidad subyacente en un sistema aparentemente tan simple como el analizado, hace referencia al hecho de que la descripción de las interacciones observadas entre un ligando y una proteína se podría pensar que puede proporcionar una visión completa y exhaustiva del sistema, más partiendo de estructuras con resoluciones atómicas. Lo que nuestros resultados indican es que esta aproximación no llega en todos los casos a establecer la suficiencia y/o necesidad de las interacciones observadas, lo cual hace intrínsecamente complejo el diseño de nuevas propiedades (sitios de unión, por ejemplo) en proteínas.

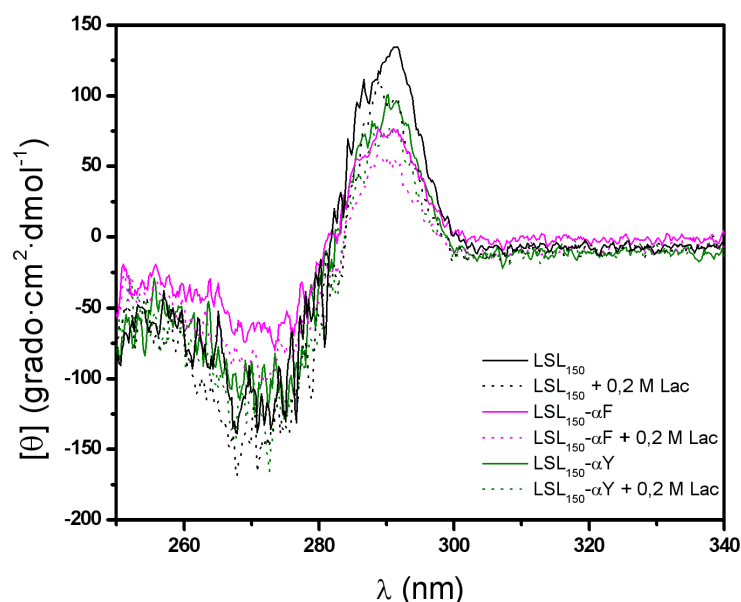
### 5.3. Estabilidad y función van en caminos opuestos

Como era de esperar por los anteriores resultados estructurales, los espectros de DC de los dos mutantes triples son esencialmente idénticos a los de la proteína silvestre en la región del UV lejano (**Figura 5.9**). Presentan un mínimo en torno a los 216 nm, que refleja el alto contenido en estructura tipo  $\beta$ . Además, la lactosa no produce ningún cambio en la estruc-

tura secundaria, ya que los espectros en presencia del ligando se superponen prácticamente a lo observado cuando la lectina está sola. En el caso de los triples mutantes en presencia de lactosa sí se observa un pequeño hombro cerca de los 230 nm. Este hombro, que ya había aparecido previamente en el espectro de DC de LBL<sub>152</sub> está relacionado con el entorno de los residuos de Trp en la estructura terciaria (Woody, 1994). Este comportamiento podría indicar que se producen pequeños cambios que afectan al microentorno de este tipo de residuos al incorporar un tercer sitio operativo de unión de azúcares.



**Figura 5.9.** Espectroscopía de diroísmo circular en el UV lejano. Comparación de los espectros de diroísmo circular de LSL<sub>150</sub> (negro) LSL<sub>150</sub>-αF (rosa) y LSL<sub>150</sub>-αY (verde) en ausencia (línea continua) y presencia (línea discontinua) de 0,2 M lactosa. Los espectros en presencia de ligando se obtuvieron dializando previamente las lectinas con tampón TSAL. Los espectros fueron medidos a 25 °C y con una cubeta de 0,1 cm de paso óptico.

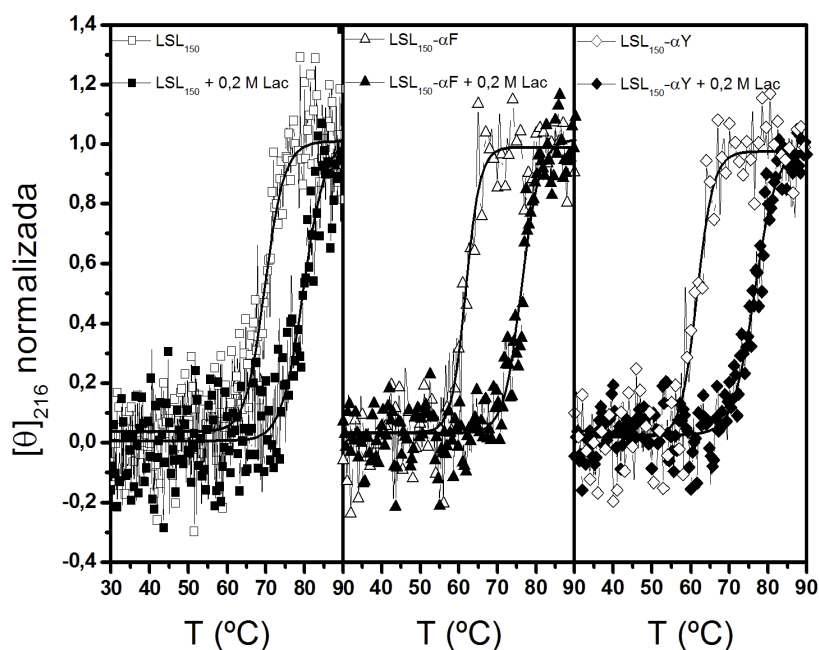


**Figura 5.10.** Espectroscopía de diroísmo circular en el UV cercano. Comparación de los espectros de diroísmo circular de LSL<sub>150</sub> (negro) LSL<sub>150</sub>-αF (rosa) y LSL<sub>150</sub>-αY (verde) en ausencia (línea continua) y presencia (línea discontinua) de 0,2 M lactosa. Los espectros en presencia de ligando se obtuvieron dializando previamente las lectinas con tampón TSAL. Los espectros fueron medidos a 25 °C y con una cubeta de 1 cm de paso óptico.

En la región del UV cercano aparece un máximo próximo a los 290 nm (Figura 5.10). A la vista de los espectros, se encuentran pequeñas diferencias que podrían indicar ambientes distintos en el entorno local de los residuos aromáticos. Estos cambios en los espectros se justificarían por la inclusión de aminoácidos aromáticos (Phe43 o Tyr43) en la

superficie de la lectina. En presencia de 0,2 M lactosa, los espectros no experimentan grandes cambios. Sí se observa, en efecto, una disminución de la intensidad del máximo, comportamiento que se repite tanto en la lectina silvestre como en los dos triples mutantes. Este hecho podría deberse a la interacción por apilamiento de la lactosa al ser reconocida por los residuos aromáticos de los sitios de unión.

Para evaluar el efecto sobre la estabilidad de la incorporación de un nuevo sitio de unión en el módulo lectina LSL<sub>150</sub>, se registró la variación de la elipticidad molar a 216 nm en función de la temperatura de la lectina silvestre y los dos triples mutantes  $\alpha$ F y  $\alpha$ Y (**Figura 5.11**). La comparación de los valores de  $T_m$  de la lectina silvestre y de los dos mutantes indica diferencias significativas en el proceso de desnaturalización. Aunque las tres especies presentan una estabilidad térmica elevada, LSL<sub>150</sub>, con una  $T_m$  de  $69,4 \pm 0,3$  °C, experimenta una disminución de la estabilidad de prácticamente 7 °C cuando se incorpora un nuevo sitio de unión de azúcares (la  $T_m$  de ambos mutantes alcanza un valor de  $\sim 62$  °C).



**Figura 5.11.** Efecto de la lactosa en la estabilidad térmica del módulo lectina N-terminal LSL<sub>150</sub> y sus respectivos mutantes. La elipticidad a 216 nm para LSL<sub>150</sub> (cuadros) LSL<sub>150</sub>-αF (triángulos) y LSL<sub>150</sub>-αY (rombos) fue monitorizada como una función de la temperatura en ausencia (símbolo vacío) y presencia (símbolo lleno) de 0,2 M lactosa. La línea continua se corresponde al ajuste a una curva sigmoidea de los datos experimentales para la determinación de la temperatura media de transición ( $T_m$ ).

Se podría sugerir que la inclusión de un nuevo sitio funcional en LSL<sub>150</sub> conlleva una pérdida considerable de estabilidad térmica. Este comportamiento ya había sido descrito con anterioridad en enzimas, en donde estabilidad y función generalmente van en caminos opuestos, ya que los aminoácidos necesarios para una catálisis óptima tienden a disminuir la estabilidad y/o a ralentizar el plegamiento (Shoichet *et al.*, 1995). De hecho, recientes

estudios con mutantes han demostrado que los sitios de unión no son óptimos para la estabilidad intrínseca de las proteínas con estructura tipo trébol  $\beta$  (Schreiber *et al.*, 1994; Shoichet *et al.*, 1995; Gosavi *et al.*, 2008).

La relación entre estabilidad y función de los tréboles  $\beta$ , por un lado, y su simetría interna, por otro, no se conoce con detalle, si bien numerosas evidencias experimentales apuntan a que la ruptura de la simetría interna propia de su estructura primaria debida a la divergencia evolutiva de los subdominios, manteniéndose la simetría tridimensional típica del trébol  $\beta$ , origina proteínas con propiedades de plegamiento más favorables (*foldable*) (Brych *et al.*, 2001). Si bien las restricciones impuestas por la simetría interna de la secuencia, en lo concerniente a residuos del núcleo hidrofóbico, son compatibles con un empaquetamiento eficaz y con una estructura estable, se observa que esta simetría está asociada con la aparición de intermedios de plegamiento, que no aparecen al relajarse dichas restricciones.

Por el contrario, la incorporación de elementos funcionales, usualmente correlacionada con una pérdida de simetría, está asociada con una ralentización del plegamiento (Gosavi *et al.*, 2008) y/o pérdida de estabilidad (Jager *et al.*, 2006). Y es que *a priori*, no hay razón alguna por la que el plegamiento, la estabilidad y la función hayan de ser todas óptimas individualmente. De hecho, por ejemplo, normalmente mutaciones en sitios de unión que conllevan pérdida de función aumentan la velocidad de plegamiento (Gruebele, 2005; Jager *et al.*, 2006).

De este modo, esperaríamos que la incorporación de un nuevo sitio de unión funcional en el subdominio  $\alpha$  provocase una pérdida de estabilidad termodinámica de la proteína o una disminución en su velocidad de plegamiento. En conclusión, las tres mutaciones en los mutantes  $\alpha F$  y  $\alpha Y$  podría explicar la pérdida de estabilidad térmica observada con respecto a la lectina silvestre LSL<sub>150</sub>.

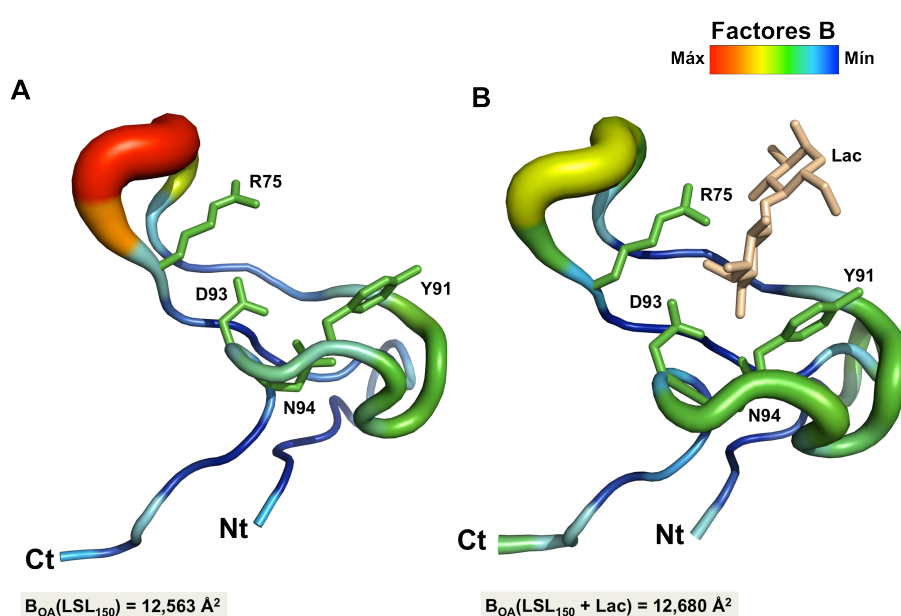
Subdominio	LSL <sub>150</sub>		LSL <sub>150</sub> :(lactosa) $\beta\gamma$	
	$B_{OA} = 12,040 \text{ \AA}^2$		$B_{OA} = 14,536 \text{ \AA}^2$	
	Factor <i>B</i> promedio ( $\text{\AA}^2$ )	$K_{\text{subdominio/lectina}}$	Factor <i>B</i> promedio ( $\text{\AA}^2$ )	$K_{\text{subdominio/lectina}}$
$\alpha$	12,309	1,02	15,408	1,06
$\beta$	12,563	1,04	12,680	0,87
$\gamma$	10,251	0,85	12,663	0,87

**Tabla 5.5.** Comparación de los factores *B* de los tres subdominios de LSL<sub>150</sub> en ausencia y presencia de lactosa. La normalización de los factores *B* se ha llevado a cabo mediante el cálculo de la razón  $K_{\text{subdominio/lectina}}$  que se obtiene como cociente entre el factor *B* promedio de cada subdominio y el factor *B* de la proteína completa ( $B_{OA}$ ).



Las curvas de desnaturalización térmica también muestran un aumento considerable de la estabilidad cuando las lectinas están en presencia de 0,2 M lactosa. Este aumento ( $\Delta T$ ) es ligeramente superior en los triples mutantes. Mientras que LSL<sub>150</sub> presenta una variación de  $\sim 12$  °C, los dos triples mutantes experimentan un  $\Delta T$  próximo a los 15 °C, lo cual podría deberse a la mayor unión de ligando por molécula de lectina por la presencia de un sitio funcional adicional.

Este efecto estabilizador del ligando puede estar relacionado con la pérdida de flexibilidad de la molécula. A la vista de las estructuras de la proteína silvestre LSL<sub>150</sub> con y sin ligando, se observa que la unión de lactosa hace más rígido el entorno del sitio de unión, puesto que se produce una disminución considerable de los factores *B* del subdominio después de la unión del ligando. Concretamente, cuando se comparan los factores *B* promedio de los tres subdominios, en el sitio  $\beta$  se experimenta una reducción del factor *B* promedio próxima al 16%, lo que evidencia la pérdida de flexibilidad (**Tabla 5.5**). La zona más flexible se encuentra en el bucle localizado en la parte opuesta a la molécula de lactosa, cerca del residuo de Arg-75, clave en la unión (**Figura 5.12**). El sitio  $\alpha$ , que no une ninguna molécula de lactosa en la proteína silvestre, no ve modificada su movilidad, por lo que el efecto estabilizador tiene un impacto local. Curiosamente, el sitio de mayor afinidad (sitio  $\gamma$ ) posee desde el principio una flexibilidad limitada y no experimenta estabilización, puesto que el factor *B* promedio ya posee un valor moderado antes de la unión del ligando, lo que podría justificar que las moléculas de lactosa se acomoden mejor en este subdominio.



**Figura 5.12.** Vista detallada del sitio de unión  $\beta$  en función de los factores *B* de cada residuo cuando LSL<sub>150</sub> se encuentra en ausencia (**A**) o presencia de lactosa (**B**). En ambas imágenes se representan los residuos implicados en la unión del ligando en verde. La molécula de lactosa se muestra en beige en la imagen **B** (derecha). Los factores *B* se distribuyen en función del intervalo que va de azul a rojo, donde el azul se corresponde a los valores más bajos y el rojo, a los más altos dentro de la estructura.

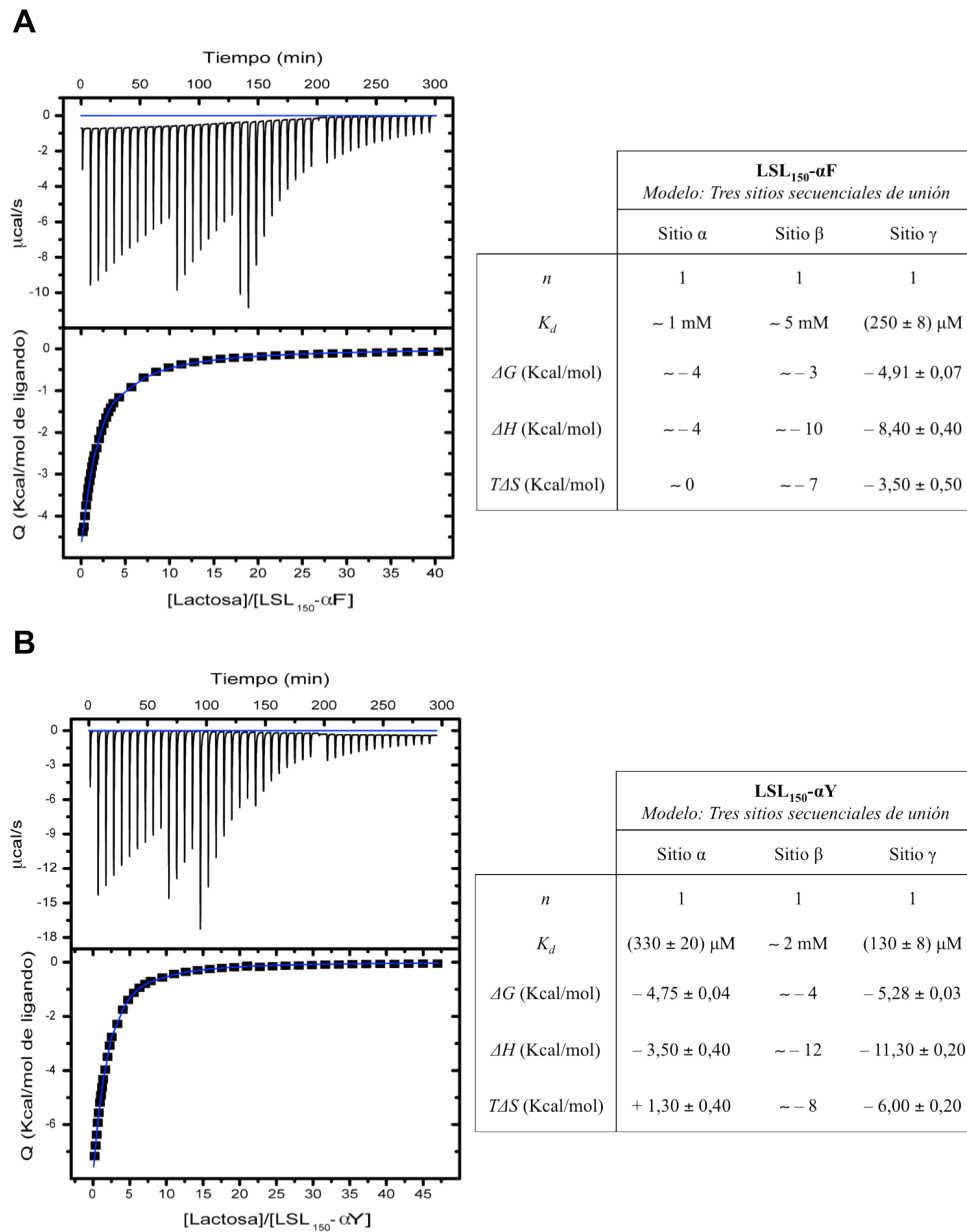


El mayor incremento en la estabilidad en los dos triples mutantes puede justificarse, en consecuencia, a un efecto similar al que tiene lugar en el sitio  $\beta$ , pero sobre los dos nuevos sitios  $\alpha$  creados. Esta hipótesis no se puede corroborar a falta de la obtención de las estructuras de los dos triples mutantes en ausencia de ligando.

#### 5.4. El nuevo sitio $\alpha$ tiene una afinidad relativamente baja

La afinidad por la lactosa de los dos triples mutantes  $\alpha F$  y  $\alpha Y$  de LSL<sub>150</sub> se estudió mediante ITC a una temperatura de 25 °C, usando una concentración final de lactosa de 60 mM. Los datos termodinámicos obtenidos no pudieron ajustarse a un modelo de un solo tipo de sitios de unión, al igual que ocurrió con los datos del módulo *wild type* de la lectina LSLa. Por esta razón, en nuestro análisis se consideraron dos modelos distintos de unión: uno, de dos tipos de sitios de unión independientes, y otro de sitios de unión de tipo secuencial. Aunque los datos termodinámicos experimentales se ajustan bien a ambos modelos, la baja afinidad de las lectinas (rango mM) (Petukhov *et al.*, 1999) hace necesario apoyar el uso de los modelos de ajuste con otras evidencias experimentales.

En primer lugar, un modelo secuencial podría aproximarnos a un contexto con sitios de unión con características estructurales diferenciadas. Los resultados de este análisis (**Figura 5.13**) revelan afinidades marcadamente diferentes para cada sitio de unión: en primer lugar, habría un sitio de elevada afinidad con una  $K_{d1}$  de  $250 \pm 30$  y  $130 \pm 8$   $\mu M$  para  $\alpha F$  y  $\alpha Y$ , respectivamente. Estos valores serían equivalentes a los observados en el sitio  $\gamma$  de la lectina silvestre, lo que implicaría que la pérdida de estabilidad global no habría afectado a la capacidad de unir ligando de este sitio de mayor afinidad relativa. En segundo lugar, habría un sitio de afinidad media con  $K_{d2}$  de  $\sim 1$  mM para LSL<sub>150</sub>- $\alpha F$  y  $330 \pm 20$   $\mu M$  para LSL<sub>150</sub>- $\alpha Y$ , que se correspondería con el nuevo sitio de unión  $\alpha$  y los valores de las constantes de disociación indicarían que el efecto de la Tyr-43 repercute en una mayor afinidad que el producido por el respectivo residuo de Phe-43. Finalmente, habría un sitio de baja afinidad, correspondiente al sitio  $\beta$ , por su similitud con el sitio  $\beta$  de LSL<sub>150</sub>, cuyas  $K_{d3}$  son de  $\sim 5$  y  $\sim 2$  mM para los mutantes  $\alpha F$  y  $\alpha Y$ , respectivamente. Este modelo secuencial implica que en primer lugar son ocupados los sitios de mayor afinidad, cuya interacción con la lactosa podría repercutir en favorecer la interacción con la lactosa en el resto de sitios, por lo que existiría cierta cooperatividad positiva. Este mecanismo ha sido propuesto para explicar la cooperatividad en algunos dominios de unión a carbohidratos de tipo ricina (Yao *et al.*, 2011), aunque hasta el momento no hay evidencias experimentales.



**Figura 5.13.** Análisis mediante ITC de la unión en solución de lactosa a los triples mutantes de LSL<sub>150</sub>: LSL<sub>150</sub>-αF (A) y LSL<sub>150</sub>-αY (B). A la izquierda se muestra la curva de valoración, así como el ajuste de los datos (azul) con el modelo de tres sitios secuenciales de unión. A la derecha se muestra una tabla con el valor de los parámetros termodinámicos determinados: la estequiometría ( $n$ ), la constante de disociación ( $K_d$ ), la energía de Gibbs ( $\Delta G$ ), la entalpía de unión ( $\Delta H$ ) y el cambio de entropía ( $T\Delta S$ ).

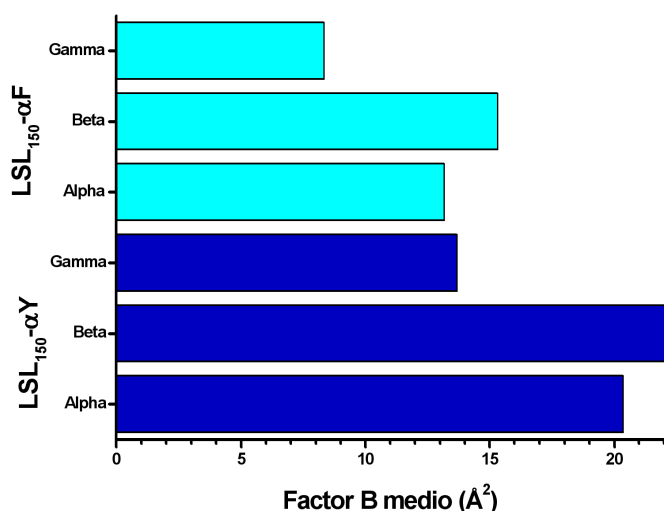
Cuando se aplica el modelo de dos tipos de sitios de unión independientes y teniendo en cuenta que los triples mutantes poseen tres sitios de unión, se asume que dos de estos sitios se están comportando de manera similar desde el punto de vista de la afinidad de la unión. En este contexto, ambos mutantes tendrían un sitio de relativa alta afinidad y dos sitios de baja afinidad [la estequiometría de los modelos se ajustó manualmente debido a que no se cumple la regla:  $[lectina] \cdot n \cdot K_d^{-1} \geq 1 - 5$  (Wiseman *et al.*, 1989)]. Las constantes de disociación ( $K_d$ ) obtenidas con este modelo para el sitio de mayor afinidad son de

540  $\pm$  40 y 879,51  $\pm$  0,02  $\mu$ M para  $\alpha$ F y  $\alpha$ Y, respectivamente (**Tabla 5.6**). Este resultado implicaría que el sitio  $\gamma$  muestra una menor afinidad que la observada en LSL<sub>150</sub> debido a la introducción de las tres mutaciones en el sitio  $\alpha$ , lo cual no se puede justificar a la luz de las estructuras, puesto que no se observan cambios en el modo de unión que se lleva a cabo en este subdominio. Por otro lado, los sitios  $\alpha$  y  $\beta$  tendrían una afinidad mucho menor, siendo la  $K_d$  de  $\sim$ 7 y  $\sim$ 30 mM para los mutantes  $\alpha$ F y  $\alpha$ Y, respectivamente. Estos resultados son difíciles de explicar desde un punto de vista estructural, lo que apoyaría la utilización del modelo secuencial para explicar los aspectos termodinámicos de los dos triples mutantes.

<i>Modelo: Dos sitios de unión independientes</i>				
	<b>LSL<sub>150</sub>-<math>\alpha</math>F</b>		<b>LSL<sub>150</sub>-<math>\alpha</math>Y</b>	
	Alta afinidad	Baja afinidad	Alta afinidad	Baja afinidad
<i>n</i>	1	2	1	2
$K_d$	540 $\pm$ 40 $\mu$ M	$\sim$ 7 mM	879, 50 $\pm$ 0,02 $\mu$ M	$\sim$ 30 mM
$\Delta G$ (Kcal/mol)	- 4,45 $\pm$ 0,05	$\sim$ - 3	- 4,17 $\pm$ 0,01	$\sim$ - 2
$\Delta H$ (Kcal/mol)	- 12,50 $\pm$ 0,60	$\sim$ - 6	- 20,40 $\pm$ 0,80	$\sim$ - 7
$T\Delta S$ (Kcal/mol)	- 8,10 $\pm$ 0,70	$\sim$ - 3	- 15,90 $\pm$ 0,80	$\sim$ - 5

**Tabla 5.6.** Análisis termodinámico de la unión a lactosa mediante ITC de los triples mutantes de LSL<sub>150</sub>: LSL<sub>150</sub>- $\alpha$ F y LSL<sub>150</sub>- $\alpha$ Y. Los diferentes parámetros termodinámicos (la estequiometría (*n*), la constante de disociación ( $K_d$ ), la energía de Gibbs ( $\Delta G$ ), la entalpía de unión ( $\Delta H$ ) y el cambio de entropía ( $T\Delta S$ )) se han determinado aplicando el modelo de dos sitios de unión independientes.

Independientemente del modelo utilizado para ajustar los datos experimentales, los ensayos de ITC demuestran que la unión a lactosa es exotérmica y está dirigida entálpicamente, resultado del establecimiento de interacciones favorables entre la lectina, los ligandos y el solvente. Si se atiende al número de interacciones observadas, éste concuerda con la afinidad respectiva de los tres sitios de unión, ya que el sitio  $\gamma$  es el que más enlaces forma. Curiosamente, el análisis de los factores *B* de los anillos de galactosa (el que presenta más contactos con los residuos de los sitios de unión) también va en la misma dirección (**Figura 5.14**). Entendiendo que factores *B* más altos implican mayor movilidad en esa zona, los sitios  $\alpha$  y  $\beta$  son los que tienen unos factores *B* más elevados. Una mayor movilidad implicaría una mayor dificultad para estabilizar la unión del ligando, lo que podría traducirse en una menor afinidad. Con todo esto, es evidente que, a pesar de la incorporación de los aminoácidos necesarios para generar un sitio tipo  $\gamma$  en el subdominio  $\alpha$  de LSL<sub>150</sub> y haber recuperado la función, obtener un grado de afinidad equivalente depende de muchos otros factores, como la estabilidad global de la proteína, su movilidad intrínseca, la ordenación del solvente y las interacciones lectina-azúcar que se lleguen a establecer.



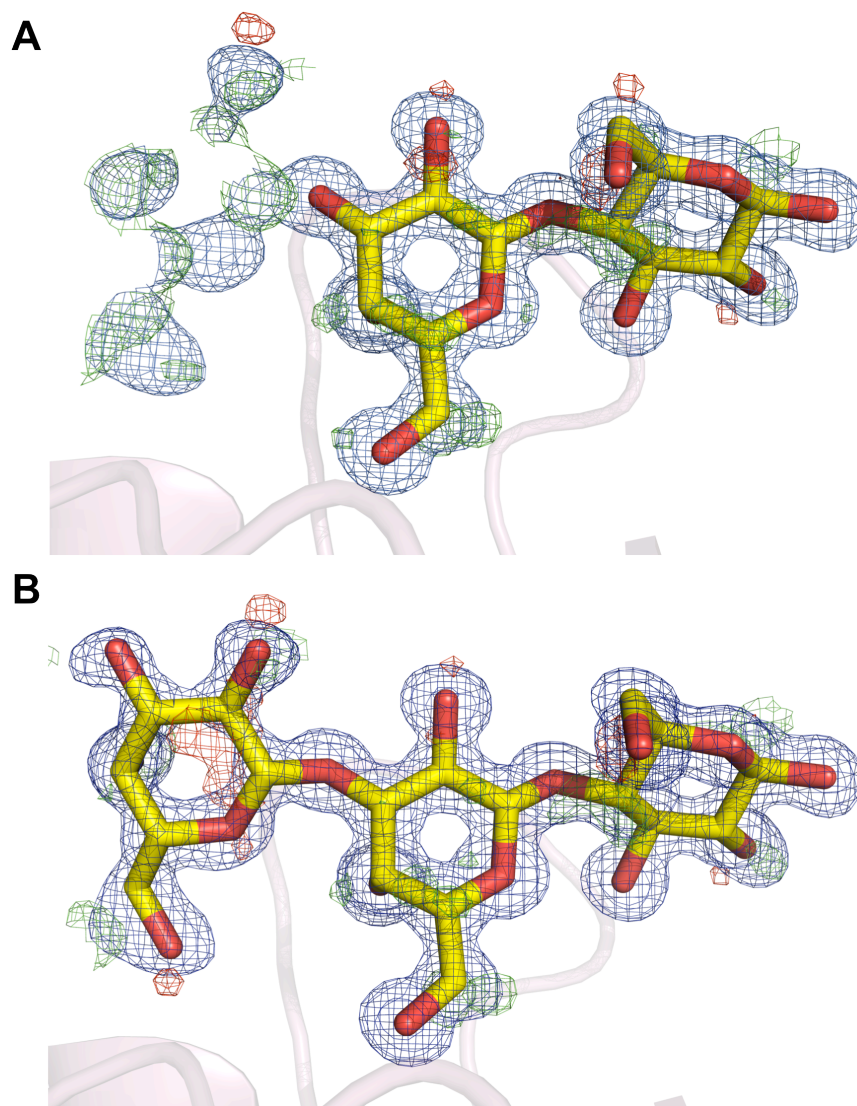
**Figura 5.14.** Análisis de los factores  $B$  medios correspondientes a los anillos de galactosa de los ligandos unidos a los sitios  $\alpha$ ,  $\beta$  y  $\gamma$  de los triples mutantes LSL<sub>150</sub>-αF (azul claro) y LSL<sub>150</sub>-αY (azul marino).

## 5.5. Identificación de un nuevo ligando en el sitio $\gamma$

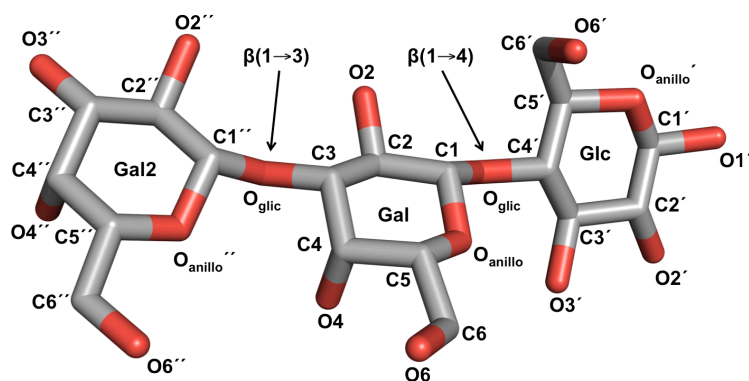
La inspección de los mapas de Fourier  $F_o - F_c$  y  $2F_o - F_c$  en el sitio de unión  $\gamma$  del triple mutante  $\alpha F$  de LSL<sub>150</sub> indicó la existencia de una densidad electrónica extra junto al oxígeno O3 de la lactosa (**Figura 5.15A**). Aparentemente, la forma adoptada por esta densidad recordaba a la de un monosacárido y no a la de moléculas de agua independientes, como suelen verse frecuentemente en los sitios de unión. Tras probar nuevos refinamientos, incorporando diferentes monosacáridos en la nueva densidad, el mejor resultado se obtuvo cuando se refinó en presencia de una unidad de  $\beta$ -D-galactopiranososa unida mediante enlace  $\beta(1 \rightarrow 3)$  a la galactosa de la molécula de lactosa (**Figura 5.15B**). El resultado obtenido demuestra, por consiguiente, que el sitio de unión  $\gamma$ , a diferencia de los otros dos sitios, es capaz de unir un trisacárido que existe en la naturaleza y que se denomina  $\beta$ -3'-galactosillactosa (GGG) (**Figura 5.16**). El análisis detallado de la densidad electrónica demuestra que este trisacárido sólo tiene un único modo de unión y que el anillo de galactosa que se reconoce preferentemente es el que está unido a la glucosa mediante enlace  $\beta(1 \rightarrow 4)$ . El análisis de los factores  $B$  de este nuevo ligando indica que el anillo 2 de galactosa o presenta más movilidad que el resto de la molécula o su ocupación es más baja. Este hecho podría justificarse si se considera que en el cristal coexisten moléculas de lectina que unen mayoritariamente lactosa junto con una pequeña fracción que une el trisacárido  $\beta$ -3'-galactosillactosa en el sitio de unión  $\gamma$ .

La procedencia de este trisacárido *a priori* es desconocida. Probablemente, sea un contaminante de la lactosa utilizada en los ensayos de cristalización. De hecho, la casa comercial (Scharlau) nos indicó que la lactosa que comercializan no es un producto 100%

puro. Generalmente, este azúcar contiene diversos contaminantes a nivel de traza, mayoritariamente derivados del azúcar de la leche como disacáridos y oligosacáridos. Este hecho coincide con que la  $\beta$ -3'-galactosillactosa se considere un galactooligosacárido (GOS) que está presente a muy baja concentración en el calostro de algunos mamíferos, como vacas y ovejas (Park, 2009). Además, este azúcar también es un producto frecuente en las reacciones de transglucosilación que llevan a cabo algunas  $\beta$ -galactosidasas como, por ejemplo, la de *Thermotoga maritima* (Marin-Navarro *et al.*, 2014). Por lo tanto, teniendo en cuenta que es posible que este compuesto esté presente en la lactosa comercial, la estructura del triple mutante nos está indicando que la afinidad del sitio de unión  $\gamma$  por la  $\beta$ -3'-galactosillactosa es muy elevada, puesto que a pesar de que se encuentra a muy baja concentración, es capaz de discriminar entre este trisacárido y la lactosa.

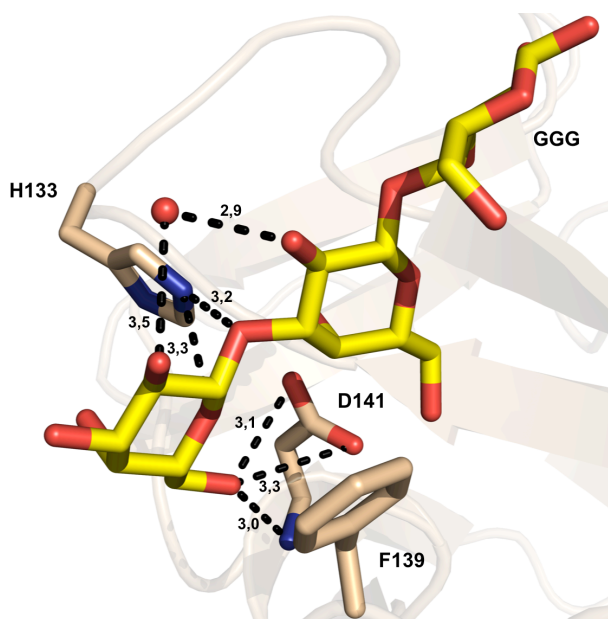


**Figura 5.15.** Densidad electrónica experimental en torno al ligando del sitio de unión  $\gamma$  del mutante LSL<sub>150</sub>- $\alpha$ F. En **A** se representa el ligando lactosa y en **B**, el ligando  $\beta$ -3'-galactosillactosa. La densidad electrónica del mapa  $2F_o - F_c$  a  $1\sigma$  se muestra en azul, mientras que la densidad del mapa  $F_o - F_c$  se muestra en verde ( $+3\sigma$ ) y en rojo ( $-3\sigma$ ).



**Figura 5.16** Estructura de la molécula de  $\beta$ -3'-galactosillactosa ([D-galactopiranosil- $\beta$ (1 $\rightarrow$ 3)]D-galactopiranosil- $\beta$ (1 $\rightarrow$ 4)-D-glucopiranosil o [Gal $\beta$ (1 $\rightarrow$ 3)]Gal $\beta$ (1 $\rightarrow$ 4)Glc) en modelo de varillas, donde se indica la numeración de todos sus átomos (carbono en gris y oxígeno en rojo).

El reconocimiento del trisacárido implica nuevas interacciones mediante enlaces de hidrógeno que se establecen con el anillo 2 de galactosa (**Figura 5.17**). Se observa en primer lugar que el átomo de oxígeno O6'' es capaz de formar hasta tres enlaces de hidrógeno con el residuo de Asp-141, dos con sendos oxígenos de la cadena lateral y uno con el nitrógeno del enlace peptídico. Por otro lado, el nitrógeno N $\epsilon$ -2 de la His133 interacciona tanto con el oxígeno del anillo como con el oxígeno del enlace glicosídico  $\beta$ (1 $\rightarrow$ 3) entre las dos unidades de galactosa. En último lugar, aparece en juego una molécula de agua que reconoce simultáneamente a los átomos de oxígeno O2 y O2'' de los anillos de galactosa. Las interacciones descritas encajan con la descripción de un sitio extendido en lectinas. El tercer anillo probablemente no presente una afinidad muy alta, pero la unión inicial del anillo 1 de galactosa favorece el establecimiento de nuevas interacciones específicas en el sitio de unión que, por tanto, dan lugar a un incremento de afinidad global por la  $\beta$ -3'-galactosillactosa con respecto a la molécula de lactosa (Weis y Drickamer, 1996).

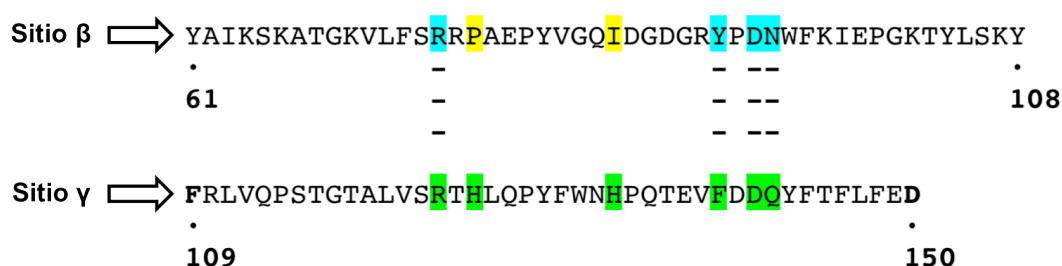


**Figura 5.17.** Vista detallada de las interacciones específicas entre el anillo de galactosa Gal2 de la  $\beta$ -3'-galactosillactosa (GGG) y los residuos del sitio de unión y del mutante LSL<sub>150</sub>- $\alpha$ F. Las moléculas de agua se representan mediante esferas de color rojo. Solo se han representado las distancias susceptibles de formar enlaces de hidrógeno dentro del intervalo de 2,0–3,5 Å. Las interacciones se representan en línea discontinua y las medidas de las distancias vienen dadas en ångströms (Å).

## 6. Rediseño del sitio $\beta$

### 6.1. Diseño, purificación, cristalización y resolución estructural

Aunque el sitio  $\beta$  dispone del residuo aromático Tyr-91, responsable del apilamiento con el anillo de galactosa, carece de las “pinzas” de histidinas responsables, entre otros factores, de la mayor afinidad de la que hace gala el sitio de unión  $\gamma$ . El resto de aminoácidos participantes en la interacción con el disacárido, no obstante, está presente en el sitio  $\beta$ , como se observa en el alineamiento de las secuencias de la **Figura 5.18**. En este sentido, con la intención de estudiar el efecto de la incorporación de las histidinas en la afinidad del sitio  $\beta$ , se diseñaron dos mutantes puntuales: Pro77His (mutante  $\beta$ 1) e Ile85His (mutante  $\beta$ 2).



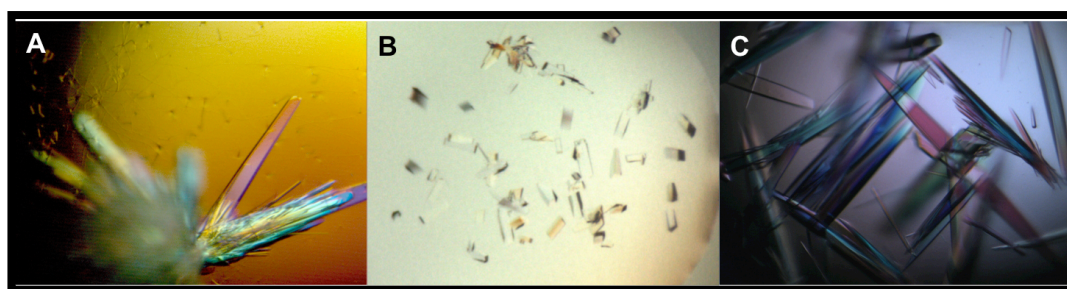
**Figura 5.18.** Alineamiento de las secuencias de aminoácidos de los subdominios  $\beta$  y  $\gamma$  del módulo lectina N-terminal LSL<sub>150</sub>. En verde se muestran los residuos involucrados en la unión de lactosa. En azul se muestran los residuos conservados en el sitio  $\alpha$  que podrían establecer contactos con el azúcar. Por último, en amarillo se muestran las posiciones seleccionadas para el diseño de mutantes con el fin de reconstituir la capacidad de unión a azúcares en el sitio  $\beta$ .

La purificación de los dos mutantes se llevó a cabo en un único paso, al igual que se había hecho con el resto de variantes. A la vista de los geles de electroforesis, ambas proteínas eluyeron con un alto nivel de pureza de la columna de Sepharose® 4B. Como cabía esperar, el rendimiento de producción fue elevado, obteniéndose 90 y 30 mg/L para los mutantes  $\beta$ 1 y  $\beta$ 2, respectivamente. El rendimiento final no se ve afectado por las mutaciones, puesto que éste se encuentra en el mismo orden de magnitud que el determinado para el resto de variantes, incluyendo el módulo lectina silvestre y los mutantes del sitio  $\alpha$ . Con el fin de eliminar la lactosa, las proteínas se dializaron exhaustivamente y se concentraron para la realización de los experimentos de cristalización.

Aunque ambos mutantes no pierden su capacidad para unir azúcares, puesto que se siguen uniendo a la resina de la cromatografía de afinidad de forma específica, sí cambian su patrón de cristalizabilidad. De hecho, del mutante  $\beta$ 1 no se consiguió ningún cristal de calidad en presencia de lactosa; sin embargo, sí se consiguieron obtener cristales en ausen-



cia de ligando tras aplicar diferentes estrategias de optimización. Finalmente, se obtuvieron cristales de calidad desde un punto de vista morfológico (**Figura 5.19A**) en la condición de cristalización 20% (p/v) PEG 6000, 0,1 M MES pH 6,0 y 0,2 M cloruro de magnesio, usando como aditivo un 15% (v/v) de glicerol en la gota con una relación 1:1:0,4 (proteína:precipitante:aditivo) y empleando una concentración de proteína de ~47 mg/mL. En cambio, por otro lado, el mutante  $\beta 2$  cristalizó sin problema tanto en su forma libre como con lactosa a una concentración de aproximadamente 45 mg/mL en dos condiciones de cristalización diferentes (**Figura 5.19B y C**): 20% (p/v) PEG 3350 y 0,2 M formiato de sodio para la lectina sola y 20% (p/v) PEG 3350, 0,1 M Bis-tris propano pH 7,5 y 0,2 M NaBr para la lectina con ligando. Los cristales crecieron al cabo de 1-2 días y su comportamiento bajo la luz polarizada indica que los cristales están muy ordenados.



**Figura 5.19.** Cristales optimizados de los mutantes sencillos del sitio  $\beta$  de LSL<sub>150</sub> obtenidos mediante difusión de vapor en gota colgante a 18 °C. **A.** Fotografía bajo luz polarizada de los cristales del mutante  $\beta 1$  (Pro77His) aparecidos en 20% (p/v) PEG 6000, 0,1 M MES pH 6,0, 0,2 M cloruro de magnesio y 15% (v/v) glicerol. **B.** Cristales del mutante  $\beta 2$  (Ile85His) aparecidos en 20% (p/v) PEG 3350 y 0,2 M formiato de sodio. **C.** Cristales del mutante  $\beta 2$  (Ile85His), observados bajo luz polarizada, aparecidos en 20% (p/v) PEG 3350, 0,1 M Bis-tris propano pH 7,5 y 0,2 M NaBr en presencia de 0,2 M lactosa.

Desafortunadamente y a pesar de la optimización de los cristales del mutante  $\beta 1$ , no se consiguieron datos de difracción de buena calidad para poder resolver la estructura tridimensional de esta lectina. No obstante, los cristales del otro mutante consiguieron difractar de forma adecuada. Los cristales del mutante  $\beta 2$  tanto en ausencia de ligando como con él fueron medidos a 100 K en las líneas ID23-1 e ID23-2 del ESRF (Grenoble, Francia), respectivamente. En ambos casos, se utilizó una longitud de onda próxima a 1 Å y se siguió una estrategia con una oscilación de 1°. El agente crioprotector que mejores resultados rindió fue el glicerol a un 25% (v/v).

A la vista de las estadísticas del procesamiento de los datos cristalográficos (**Tabla 5.7**), las resoluciones finales alcanzadas fueron elevadas, siendo de 1,65 Å para el estado libre del mutante  $\beta 2$  y 1,77 Å para el complejo con lactosa. Los cristales de la lectina libre son ortorrómbicos, pertenecen al grupo espacial  $P2_12_12_1$  y tienen una sola molécula en la



unidad asimétrica. Los cristales del complejo, en cambio, son monoclinicos (grupo espacial  $P2_1$ ) y tienen ocho moléculas en la unidad asimétrica. La estructura cristalográfica se obtuvo en los dos casos mediante reemplazo molecular, usando la estructura de LSL<sub>150</sub> como modelo. El refinamiento posterior incluyó, entre otros, refinamiento TLS y NCS para el complejo con lactosa. Los modelos finales se obtuvieron con unos valores de  $R_{work}/R_{free}$  de 0,174/0,199 y 0,167/0,202.

Parámetros	LSL <sub>150</sub> -β2	LSL <sub>150</sub> -β2:lactosa
Recogida de datos y procesamiento		
Grupo espacial	$P2_12_12_1$	$P2_1$
Moléculas en la UA	1	8
Celda unidad (Å)	34,88, 59,90, 66,89	36,56, 142,50, 114,40
Longitud de onda (Å)	1,067	0,873
Resolución (Å)	35–1,65	29–1,77
Reflexiones totales	112.761	455.838
Reflexiones únicas	16.853	109.972
Complejidad	99,3 (100)	98,8 (97,1)
Multiplicidad	6,5 (6,7)	4,1 (4,1)
$R_{merge}$ (%)	8,5 (39,9)	10,8 (47,0)
$I/\sigma(I)$	13,5 (5,1)	9,9 (3,2)
Refinamiento		
$R_{work}/R_{free}$	0,177/0,199	0,167/0,202
RMSD de los enlaces (Å)	0,006	0,007
RMSD de los ángulos (°)	1,191	1,237
Factor $B$ medio (Å <sup>2</sup> )	26,42	21,15
Número de átomos		
Proteína	1.182	9.688
Glicerol	0	6
Lactosa	0	368
Cloruro	0	4
Agua	113	792
Mapa de Ramachandran		
Favorecidos (%)	97,9	96,5
Permitidos (%)	2,1	2,9
Prohibidos (%)	0	0,6

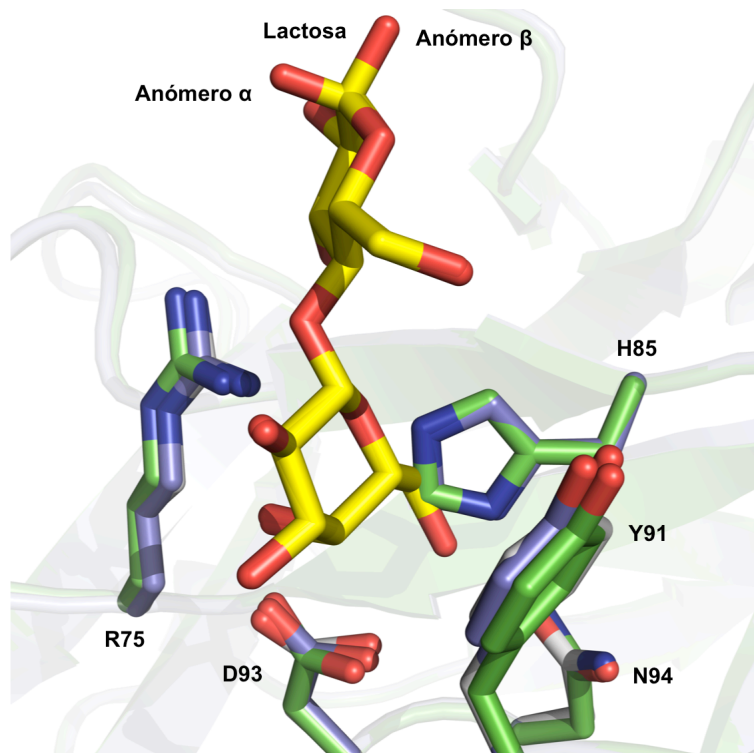
\*Los valores de la concha de mayor resolución se muestran en paréntesis.

**Tabla 5.7.** Datos de difracción y estadísticas del refinamiento de las estructuras del mutantes β2 (Ile85His) de LSL<sub>150</sub> en ausencia y presencia de 0,2 M lactosa.

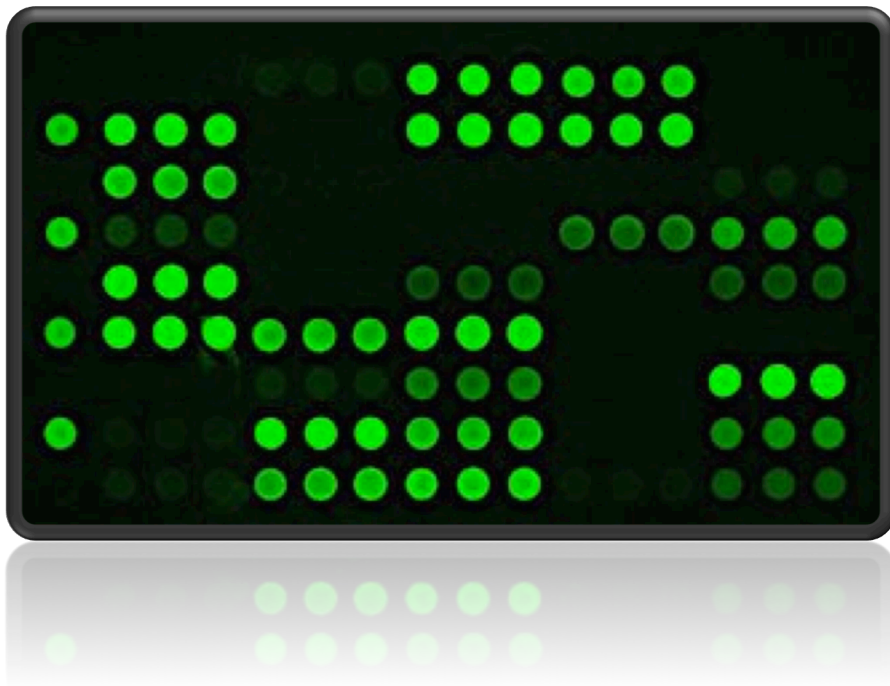
En comparación con la estructura de la lectina silvestre, los mutantes conservan la misma estructura global. Los RMSD de los Cα son de 0,34 y 0,63 para las estructuras libre y en complejo, respectivamente, lo que supone que la mutación puntual Ile85His no provoca ningún cambio estructural de relevancia.

## 6.2. Efecto de la mutación Ile85His en la unión de lactosa

Sorprendentemente, la inspección detallada del sitio  $\beta$  en las estructuras con y sin lactosa de LSL<sub>150</sub>- $\beta$ 2 reveló que éste ha perdido su capacidad de unión de azúcares. Aunque la disposición espacial de los residuos que intervienen en la unión (Arg-75, Tyr-91, Asp-93 y Asn-94) no varía con respecto a la proteína silvestre, la orientación del nuevo residuo de His-85 impide la acomodación correcta del azúcar (**Figura 5.20**). La cadena lateral de la His-85 se interpondría entre el aromático y el ligando, ocupando la zona que debería albergar a los átomos C6 y O6 del anillo de galactosa. Este factor es crítico para impedir el reconocimiento y hacer que el sitio  $\beta$  deje de ser funcional. En este sentido, a pesar de que a nivel de la secuencia las pinzas de histidinas representarían el elemento clave para un aumento de la afinidad, su disposición espacial no es equivalente a la que adopta la His-133 del sitio  $\gamma$  y, por lo tanto, no ejerce la misma función en el reconocimiento.



**Figura 5.20.** Vista detallada del sitio  $\beta$  del módulo lectina LSL<sub>150</sub> y del respectivo mutante LSL<sub>150</sub>- $\beta$ 2 (Ile85His). La molécula de lactosa (amarillo) así como los residuos partícipes de la unión del disacárido y el residuo mutado se muestran con el modelo de varillas. Los residuos de la proteína silvestre se representan en gris; los del mutante en su estado libre, en verde; y los del mutante en el complejo, en azul violáceo.



## RESULTADOS Y DISCUSIÓN

Análisis funcional de  
la unión a azúcares



## VI

## RESULTADOS Y DISCUSIÓN

## Análisis funcional de la unión a azúcares

7. Estudio de la especificidad de LSL<sub>150</sub> y LBL<sub>152</sub>

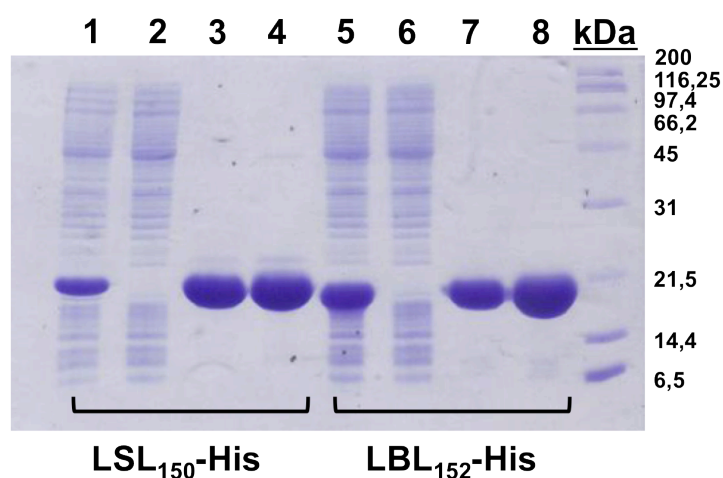
Desde finales del siglo XX, los azúcares están en el punto de mira de muchas investigaciones en Biología y Biomedicina (Park *et al.*, 2013). Esto se debe sobre todo a los últimos descubrimientos en los que se ha descrito la importancia de los azúcares en multitud de procesos biológicos (Varki, 1993). Todos ellos tienen en común la interacción de un azúcar con al menos una proteína normalmente conocida como GBP (*glycan binding protein*). Esta interacción tiene un sentido biológico en tanto que implica una transferencia de información en un contexto biológico determinado que es relevante para la célula, razón por la cual es crucial conocer detalladamente la naturaleza de la misma. Siguiendo con este razonamiento, llevamos a cabo estudios sobre la especificidad de la unión de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> con el fin de encontrar pistas acerca del rol que podrían desempeñar estas lectinas en su entorno natural. Hasta ahora no hay información disponible de los ligandos biológicos de ambas lectinas; simplemente se conoce su especificidad por  $\beta$ -galactósidos desde el punto de vista de la experimentación *in vitro*. Esta aproximación ayudará a comprender mejor la complejidad de los procesos de reconocimiento que podrían tener lugar *in vivo*, tanto en lo que a la estructura como al modo de unión del ligando se refiere. Con esta intención, se realizaron diferentes *microarrays*, que permiten hacer un barrido amplio de potenciales ligandos rápidamente y consumiendo muy poca cantidad de muestra.

**7.1. LSL<sub>150</sub> y LBL<sub>152</sub> presentan un perfil de especificidad por azúcares similar.**

7.1.1. Los módulos lectina se producen normalmente con una etiqueta de seis histidinas

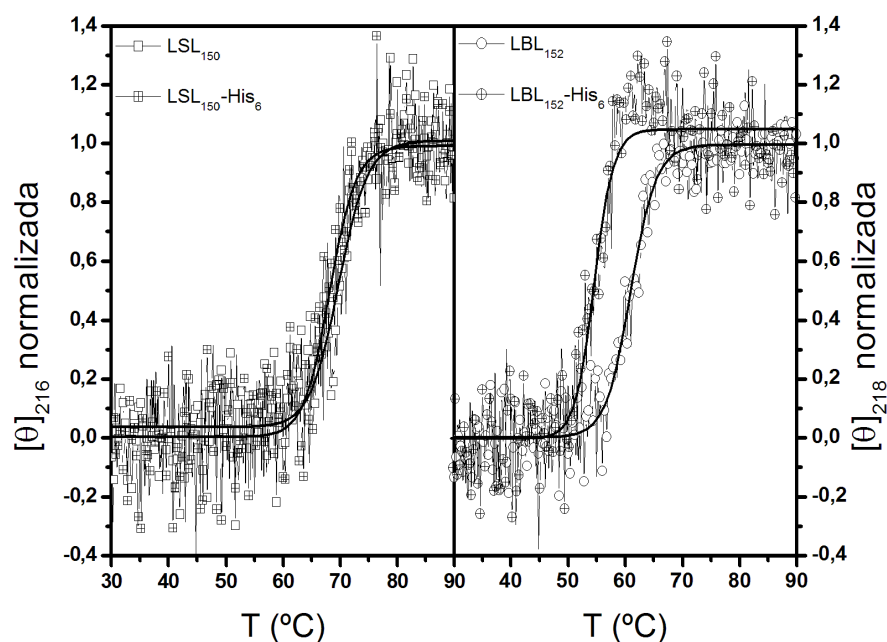
Existen varias aproximaciones experimentales para poder identificar qué azúcares inmovilizados en un *glycan array* interaccionan con una determinada lectina: marcaje directo con fluoróforos, biotinilación, uso de anticuerpos en tándem, etc. (Park *et al.*, 2013). En el caso de las lectinas LSL<sub>150</sub> y LBL<sub>152</sub>, se optó por la producción de sendos módulos lectina con una etiqueta de seis histidinas (His<sub>6</sub>) en el extremo C-terminal. Esta etiqueta de afinidad es reconocida específicamente por anticuerpos monoclonales anti-*His tag* ( $\alpha$ -His) marcados con una sonda fluorescente. Hemos considerado este procedimiento puesto que el extremo C-terminal de ambas lectinas está localizado en el lado opuesto a los sitios de unión de azúcares (**Figuras 4.5 y 4.27**) por lo que su presencia no debería interferir en la interacción con los azúcares inmovilizados. Además, la etiqueta de afinidad permite llevar a cabo una estrategia de purificación libre del uso de azúcares, que pudieran permanecer unidos a las lectinas, interfiriendo por ello en los posteriores ensayos.

Las construcciones LSL<sub>150</sub>-His y LBL<sub>152</sub>-His se expresaron en células de *E. coli* BL21(DE3) y se purificaron mediante cromatografía de afinidad de unión a Ni<sup>2+</sup> seguida de una cromatografía de exclusión molecular, tal y como se describe en la sección 5.1 de *Materiales y Métodos*. A la vista del gel de electroforesis de la **Figura 6.1**, las proteínas se producen con un nivel alto de pureza y en gran cantidad, como se desprende de la intensidad de las bandas. De hecho, la incorporación de la etiqueta de afinidad no repercute en absoluto en el rendimiento de producción, que sigue manteniéndose elevado. Concretamente, se produjeron 66 mg/L de LSL<sub>150</sub>-His y 99 mg/L de LBL<sub>152</sub>-His, lo que indica que la alta tasa de traducción de los correspondientes ARNm no se ve alterada por una extensión de la secuencia en el extremo C-terminal.

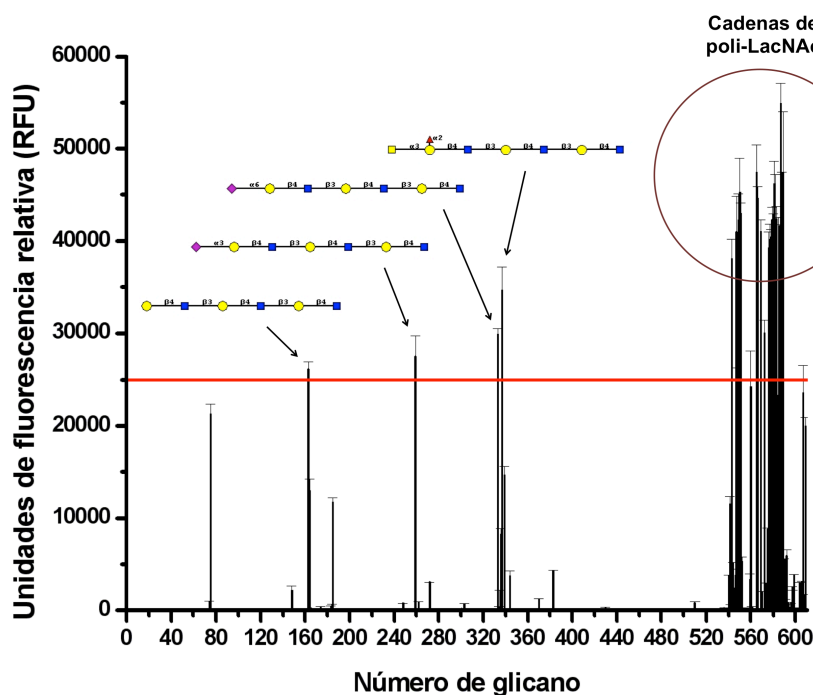


**Figura 6.1.** SDS-PAGE (15% de poliacrilamida) de la purificación de las construcciones LSL<sub>150</sub>-His y LBL<sub>152</sub>-His. Los carriles del 1-4 pertenecen a la purificación de LSL<sub>150</sub>-His y los carriles del 5-8, a la de LBL<sub>152</sub>-His. Los carriles 1 y 5 se corresponden con las fracciones solubles (FS) de los extractos celulares; los carriles 2 y 6, con las fracciones no retenidas de la columna HisTrap HP (FT); los carriles 3 y 7, con las fracciones eluidas (FE1) de la cromatografía de afinidad a Ni<sup>2+</sup>; y, por último, los carriles 4 y 8, con las fracciones eluidas (FE2) de la cromatografía de exclusión molecular.

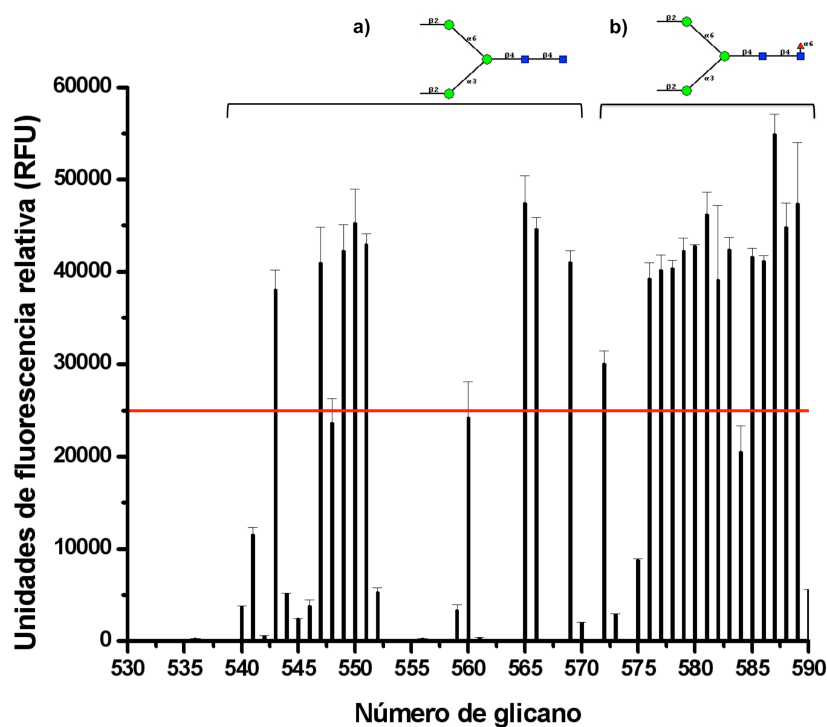
Antes de llevar a cabo el ensayo de especificidad, se realizó una prueba para comprobar que los sitios de unión no se habían visto alterados en las construcciones con la cola de histidinas. Ambas lectinas siguen quedando retenidas en columnas de Sepharose® 4B y se eluyen con facilidad al pasar una solución con 0,2 M lactosa a través de ellas. Este resultado demuestra que la etiqueta de afinidad no afecta a la capacidad de unión de las lectinas, por lo que no supone un inconveniente para llevar a cabo el *glycan array*. Además de este ensayo, se comprobó si la estabilidad térmica de ambas lectinas se mantenía a pesar de incorporar una cola de seis histidinas. Mediante dicroísmo circular, se determinaron las temperaturas de transición medias. Los resultados (**Figura 6.2**) demuestran que la temperatura de transición para LSL<sub>150</sub>-His es de  $68,2 \pm 0,3$  °C, apenas un grado centígrado más baja que la determinada para el módulo lectina solo. Esta diferencia no supone un cambio en la estabilidad térmica significativo. En cambio, la  $T_m$  correspondiente a LBL<sub>152</sub>-His es de  $54,6$  °C  $\pm$   $0,2$  °C, es decir, unos 6 °C menor que la observada para LBL<sub>152</sub>. Aún así, la estabilidad sigue siendo lo suficientemente alta para poder llevar a cabo ensayos a temperatura ambiente, como el *glycan array*, sin problema.



**Figura 6.2.** Efecto de la incorporación de una cola de seis histidinas en el extremo C-terminal en la estabilidad térmica de los módulos lectina N-terminales LSL<sub>150</sub> y LBL<sub>152</sub>. La elipticidad a 216 nm para LSL<sub>150</sub> (cuadros) y a 218 nm para LBL<sub>152</sub> (círculos) fue monitorizada como una función de la temperatura en ausencia (símbolo vacío) y presencia (símbolo con una cruz) de la etiqueta de afinidad. La línea continua se corresponde al ajuste a una curva sigmoidea de los datos experimentales para la determinación de la temperatura media de transición ( $T_m$ ).

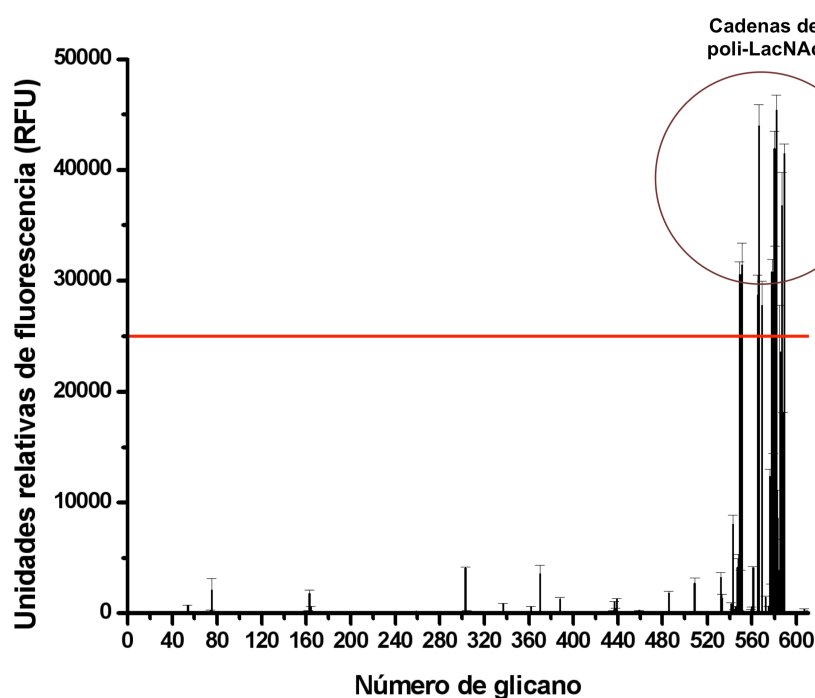
**A****B**

Cadenas de poli-LacNAc lineales y ramificadas con estructuras tipo:

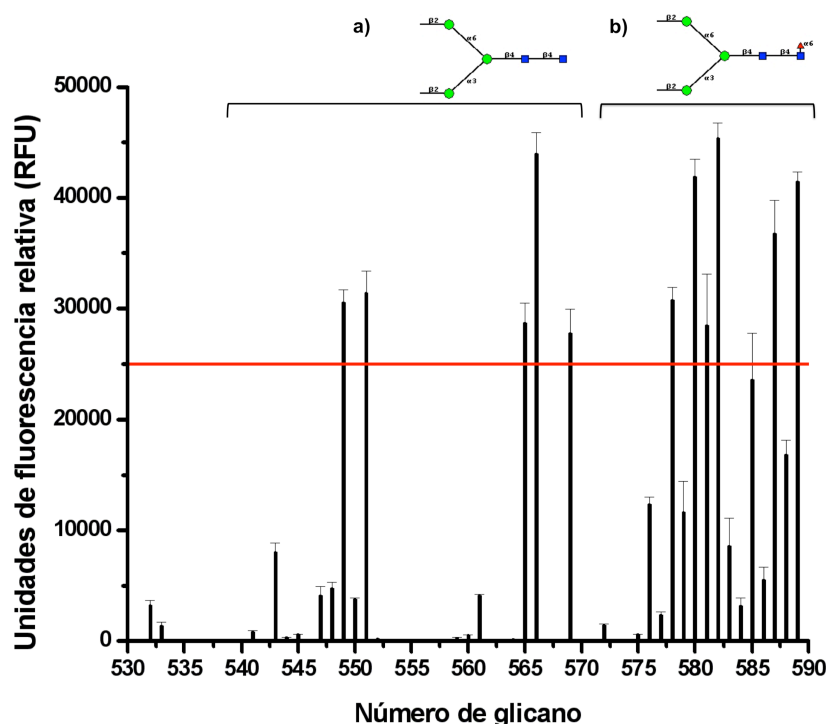


**Figura 6.3.** Especificidad de unión a carbohidratos de LSL<sub>150</sub>-His. La unión de LSL<sub>150</sub>-His a la versión 5.0 del *Mammalian Printed Array* fue analizada por el *Consortium for Functional Glycomics* (CFG). Los resultados mostrados se corresponden a la media de medidas de fluorescencia tomadas por triplicado para la lectina a una concentración de 2 µg/mL. La *línea roja* (25.000 RFU) separa los valores de intensidad más altos del resto, indicando qué azúcares se unen con mayor afinidad. En **A** se muestran los resultados frente a los 611 glicanos que comprenden el *array*. En **B** se muestra una imagen ampliada de la zona en la que hay mayor acumulación de resultados positivos.



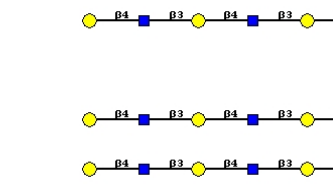
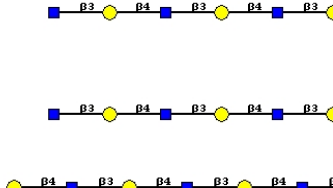
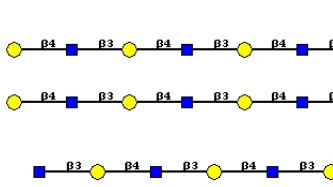
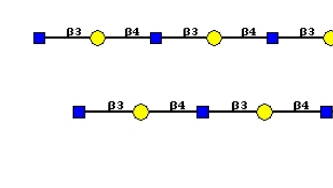
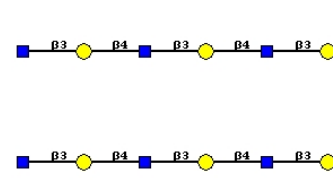
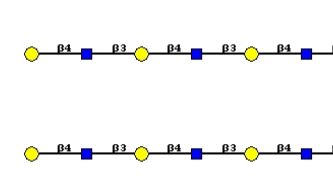
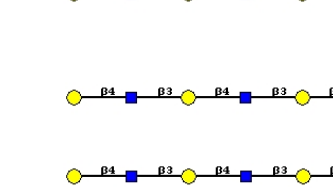
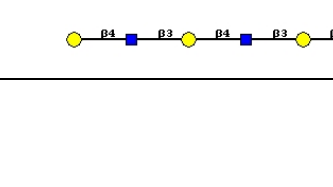
**A****B**

Cadenas de poli-LacNAc lineales y ramificadas con estructuras tipo:

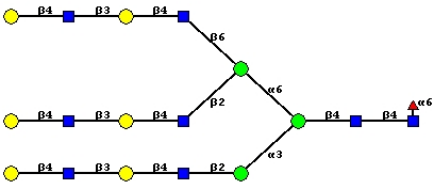
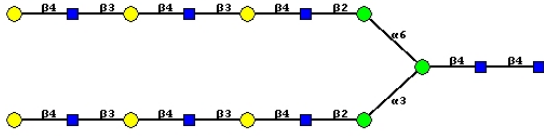
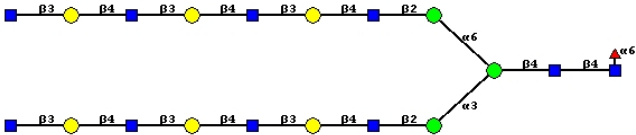


**Figura 6.4.** Especificidad de unión a carbohidratos de LBL<sub>152</sub>-His. La unión de LBL<sub>152</sub>-His a la versión 5.0 del *Mammalian Printed Array* fue analizada por el Consortium for Functional Glycomics (CFG). Los resultados mostrados se corresponden a la media de medidas de fluorescencia tomadas por triplicado para la lectina a una concentración de 2  $\mu\text{g/mL}$ . La línea roja (25.000 RFU) separa los valores de intensidad más altos del resto, indicando qué azúcares se unen con mayor afinidad. En **A** se muestran los resultados frente a los 611 glicanos que comprenden el array. En **B** se muestra una imagen ampliada de la zona en la que hay mayor acumulación de resultados positivos.

**Tabla 6.1-I.** Glicanos reconocidos en el *glycan array* por LSL<sub>150</sub>-His

Número de glicano	Glicano	Intensidad media (RFU)
587		54.851
565		47.400
589		47.326
581		46.137
550		45.219
588		44.734
566		44.593
551		42.949
580		42.708

**Tabla 6.1-II.** Glicanos reconocidos en el *glycan array* por LSL<sub>150</sub>-His

Número de glicano	Glicano	Intensidad media (RFU)
583		42.344
549		42.258
579		42.224

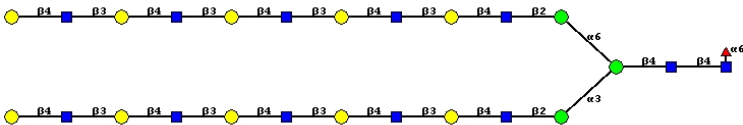
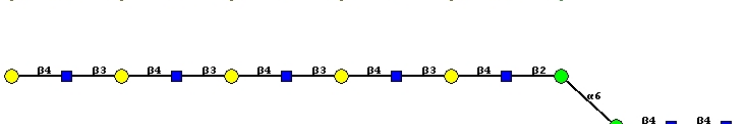
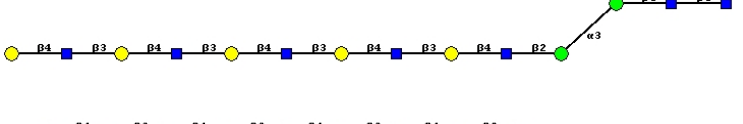
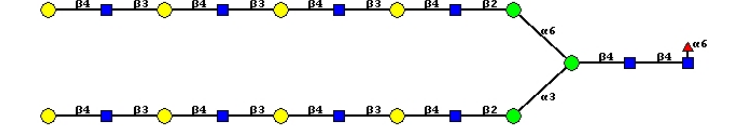
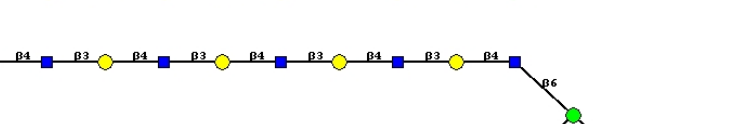
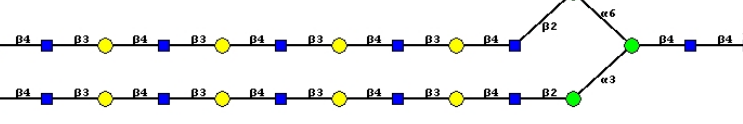
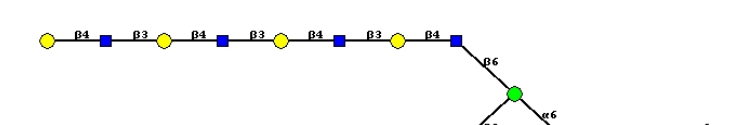
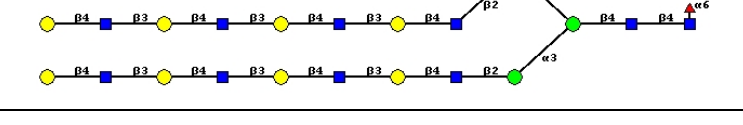

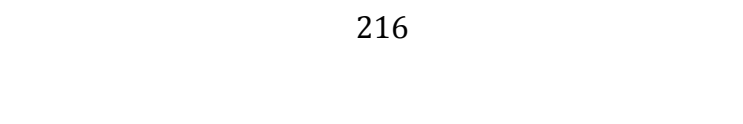
**Tabla 6.1.I-II.** Estructura de los doce glicanos que reconoce LSL<sub>150</sub>-His con mayor afinidad. Se muestra el número del glicano en la versión 5.0 del *Mammalian Printed Array* (CFG) junto con su estructura química y ordenados según la intensidad relativa de fluorescencia media (RFU) obtenida en el ensayo realizado a una concentración de 2 µg/mL de LSL<sub>150</sub>-His.

### 7.1.2. LSL<sub>150</sub> y LBL<sub>152</sub> tienen predilección por oligosacáridos ricos en poli-LacNAc

El estudio de la especificidad de LSL<sub>150</sub>-His y LBL<sub>152</sub>-His demostró que ambas proteínas presentan, a grandes rasgos, un patrón similar de reconocimiento de azúcares (**Figuras 6.3 y 6.4**). En los dos casos, se observa que los oligosacáridos que muestran una mejor unión son aquellos constituidos por repeticiones del disacárido N-acetilactosamina (LacNAc), denominados comúnmente como poli-LacNAc (**Tablas 6.1.I-II y 6.2.I-II**). Estos glicanos lineales de poli-LacNAc surgen de la unión mediante enlaces glicosídicos de tipo  $\beta(1\rightarrow3)$  de unidades de LacNAc, las cuales a su vez se forman mediante el establecimiento de un enlace  $\beta(1\rightarrow4)$  entre los residuos de galactosa (Gal) y N-acetilglucosamina (GlcNAc). Los resultados del *glycan array*, además, indican claramente que los dos módulos lectina son capaces de llevar a cabo una unión multivalente (actuando con más de un sitio de unión al mismo tiempo), puesto que la intensidad de fluorescencia detectada aumenta drásticamente cuando los oligosacáridos poseen ramificaciones (glicanos de tipo tanto biantenarico como triantenarico). Por otro lado, la afinidad es mayor a medida que aumenta

la longitud de la cadena, esto es, con la incorporación de más unidades de LacNAc en la molécula. De hecho, la mayor unión se observa en azúcares biantenarios en los que ambas cadenas tienen cinco unidades de LacNAc consecutivas. Estos resultados sugieren una correlación entre el número de residuos y el número de contactos que pueden establecerse entre la proteína y el ligando, y por ello con la afinidad relativa. Considerando que este comportamiento frente a la presencia de ramificaciones y a la longitud de los azúcares se reproduce con ambos módulos lectina, estas características estructurales de los azúcares no determinan (con este conjunto de azúcares empleado) diferencias de la especificidad entre ellas. Cabe destacar que el perfil de unión a las cadenas de poli-LacNAc no es dependiente de la concentración de lectina utilizada en el ensayo, puesto que a concentraciones bajas (0,02  $\mu\text{g/mL}$ ) se observa el mismo patrón.

**Tabla 6.2-I.** Glicanos reconocidos en el *glycan array* por LBL<sub>152</sub>-His

Número de glicano	Glicano	Intensidad media (RFU)
582		45.322
		
566		43.945
		
580		41.821
		
589		41.405
		
587		36.718
		

**Tabla 6.2-II.** Glicanos reconocidos en el *glycan array* por LBL<sub>152</sub>-His

Número de glicano	Glicano	Intensidad media (RFU)
551		31.365
578		30.697
549		30.496
565		28.609
581		28.410
569		27.710
585		23.465

**Tabla 6.2.I-II.** Estructura de los doce glicanos que reconoce LBL<sub>152</sub>-His con mayor afinidad. Se muestra el número del glicano en la versión 5.0 del *Mammalian Printed Array* (CFG) junto con su estructura química y ordenados según la intensidad relativa de fluorescencia media (RFU) obtenida en el ensayo realizado a una concentración de 2 µg/mL de LBL<sub>152</sub>-His.

Con respecto a LBL<sub>152</sub>-His, la unión disminuye a medida que las cadenas de poli-LacNAc incorporan algún tipo de modificación. Por ejemplo, la inclusión de residuos de ácido siálico o fucosa se traduce en una menor unión. Asimismo, aunque reconoce muy

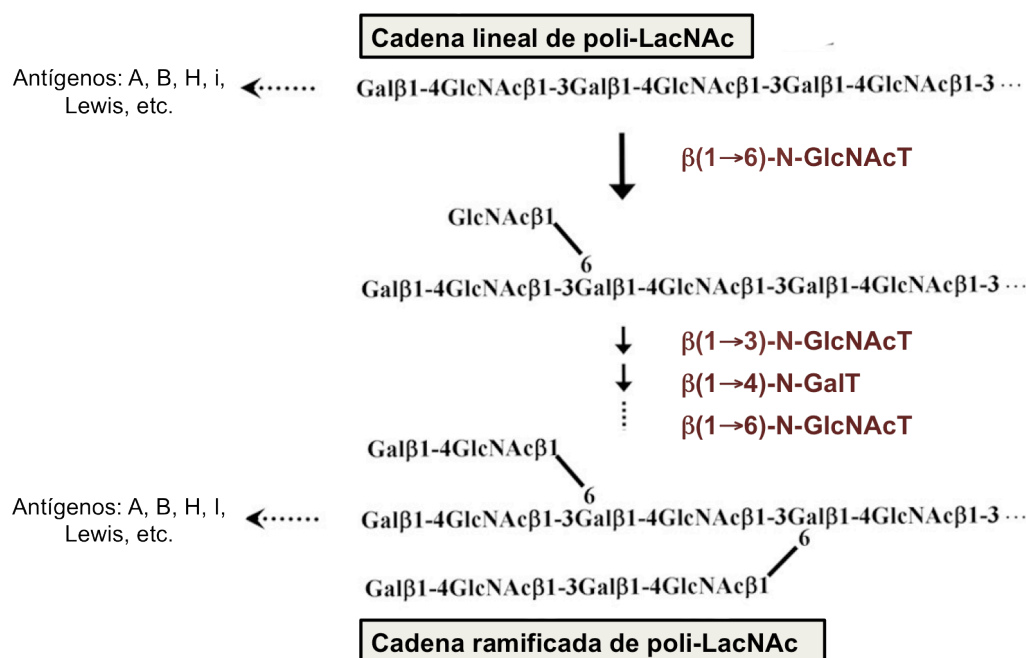
bien cadenas con galactosa o N-acetilglucosamina (GlcNAc) en posición terminal, la unión es bastante más baja en el segundo caso. Esto podría significar que la acomodación de los residuos de tipo galactosa está más favorecida que los de tipo glucosa, lo que indicaría que la galactosa es el principal motivo que confiere la especificidad de esta lectina. En último lugar, a diferencia de lo que ocurre con LSL<sub>150</sub>-His, el módulo lectina de *L. bicolor* reconoce relativamente mejor azúcares cortos de tipo bi, tri y tetraantenario, constituidos únicamente por una unidad de LacNAc en cada cadena. Se ha observado que este comportamiento, en cambio, es fuertemente dependiente de la concentración de proteína, aumentando la intensidad de fluorescencia cuanto más concentración de proteína se utiliza en el ensayo, lo que podría estar dando una idea de la afinidad.

Por otro lado, LSL<sub>150</sub>-His presenta un espectro de unión más amplio, reconociendo más azúcares. Aunque los azúcares de tipo poli-LacNAc son los que dan lugar a una señal más intensa (**Tabla 6.1.I-II**), estos pueden incorporar modificaciones, sobre todo en los extremos terminales, sin que ello signifique una caída brusca de la unión (**Figura 6.3A**). Así, la unión apenas se ve alterada por la presencia de un residuo de GlcNAc en posición terminal, lo que supone un comportamiento completamente diferente al observado en el caso de LBL<sub>152</sub>-His. Además, se observa que existen uniones relativamente elevadas cuando la galactosa terminal presenta algún tipo de modificación. Por ejemplo, la incorporación de un residuo de ácido siálico (Neu5Ac), unido mediante un enlace  $\alpha$ -glicosídico en posición (2→3) ó (2→6) a la galactosa terminal en cadenas de poli-LacNAc con dos y tres unidades de LacNAc, no afecta a la unión e indica un comportamiento distintivo con respecto al otro módulo lectina. Finalmente, también cabe destacar la elevada unión que muestra por carbohidratos lineales cortos (dos o tres unidades de LacNAc) que contienen en posición terminal el determinante GalNAc $\alpha$ (1→3)[Fuc $\alpha$ (1→2)]Gal $\beta$ -, característico del grupo sanguíneo humano A. Este resultado justifica la capacidad que la lectina LSLa tiene para aglutinar y lisar eritrocitos humanos de tipo A, como ya se había descrito previamente (Tateno y Goldstein, 2003).

Los 611 azúcares testados en el *glycan array* se corresponden con la versión 5.0 del *Mammalian Printed Array* y pueden encontrarse junto con sus nombres sistemáticos y estructuras químicas en la página web del *Consortium for Functional Glycomics* (CFG; [www.functionalglycomics.org](http://www.functionalglycomics.org)). En esta misma página web también están accesibles los datos de los ensayos de unión de los dos módulos lectina a las diferentes concentraciones.

### 7.1.3. Las estructuras ricas en LacNAc tienen relevancia biológica

Como se ha visto en el apartado anterior, el perfil de unión tanto de LSL<sub>150</sub>-His como LBL<sub>152</sub>-His indica una elevada especificidad hacia azúcares que contienen unidades de LacNAc en su interior. Este tipo de estructuras son muy frecuentes en la naturaleza y aparecen en muchos de los glicanos incorporados por N-glicosilación en las glicoproteínas de organismos vertebrados (van den Berg *et al.*, 2004). La formación de estructuras de tipo poli-LacNAc se debe principalmente a la acción conjunta de las enzimas  $\beta(1\rightarrow3)$ -N-acetilglucosaminiltransferasa y  $\beta(1\rightarrow4)$ -galactosiltransferasa. No obstante, las cadenas de poli-LacNAc lineales pueden volverse ramificadas gracias a la participación de otra enzima, la  $\beta(1\rightarrow6)$ -N-acetilglucosaminiltransferasa, que cataliza la formación de enlaces glicosídicos de tipo  $\beta(1\rightarrow6)$ , que contribuyen, por tanto, a aumentar la complejidad de los glicanos (Piller *et al.*, 1984) (**Figura 6.5**). La aparición tan extendida de estructuras de tipo poli-LacNAc en la naturaleza se debe a que juegan un papel biológico de relevancia en muchos organismos, como los seres humanos. De hecho, se ha descrito hasta ahora que son importantes en el tráfico celular, en procesos de adhesión, en el recambio de receptores celulares e incluso en el reconocimiento de microorganismos patógenos (Salomonsson *et al.*, 2010).



**Figura 6.5.** Estructuras lineales y ramificadas de cadenas de poli-LacNAc. Varios glicoepítopos sanguíneos se construyen a partir de cadenas de poli-LacNAc y la participación sucesiva de diferentes enzimas con actividad transferasa:  $\beta(1\rightarrow3)$ -N-acetilglucosaminiltransferasa,  $\beta(1\rightarrow4)$ -galactosiltransferasa y  $\beta(1\rightarrow6)$ -N-acetilglucosaminiltransferasa. Entre los glicoepítopos más característicos, se encuentran los siguientes: grupo sanguíneo A, grupo sanguíneo B, antígeno H, antígenos i e I y antígenos de la serie Lewis ( $\text{Le}^a$ ,  $\text{Le}^b$ ,  $\text{Le}^x$ ,  $\text{Le}^y$ ,  $\text{sLe}^a$  y  $\text{sLe}^x$ ).

Quizás, los azúcares de tipo poli-LacNAc sean más conocidos por constituir el núcleo sobre el que se asientan la mayor parte de glicoepítomos que conforman el sistema sanguíneo ABO(H) y Lewis (Kupper *et al.*, 2012). Cada uno de estos glicoepítomos se diferencia del resto por modificaciones que suelen tener lugar sobre los residuos terminales de las cadenas de poli-LacNAc. En este sentido, es destacable el caso de los antígenos i e I de la sangre que están constituidos única y exclusivamente por cadenas de poli-LacNAc. Estos antígenos no solo aparecen en los eritrocitos, sino que están distribuidos en muchos otros tejidos. La principal diferencia entre los antígenos de tipo i con los de tipo I es la ramificación del glicano, que a su vez es dependiente del estado de desarrollo del organismo. Así, en fetos y neonatos predomina el antígeno i (cadenas lineales), que progresivamente va aumentando su complejidad con la adición de ramificaciones, por lo que en la edad adulta, en cambio, el antígeno I (cadenas ramificadas) es el que se expresa mayoritariamente (Twu *et al.*, 2010). Mientras que esta progresión es normal y juega un papel significativo en la diferenciación de los eritrocitos, una evolución inversa (de I a i) estaría relacionada con desórdenes hematológicos, tales como talasemias, anemia hipoplásica y leucemia (Giblett y Crookston, 1964). Debido a la capacidad de la lectinas LSL<sub>150</sub>-His y LBL<sub>152</sub>-His para unir con mayor afinidad estructuras de tipo poli-LacNAc ramificadas, su utilización en diagnóstico para la identificación de este tipo de desórdenes podría considerarse de interés.

Debido a su implicación en multitud de procesos biológicos diferentes, las cadenas de poli-LacNAc se estudian con el fin de comprender mejor su función fisiológica, así como con la intención de desarrollar posibles marcadores de diagnóstico y agentes terapéuticos (Paulson *et al.*, 2006). Por ejemplo, el antígeno A de la sangre, reconocido por LSL<sub>150</sub>-His, como se ha visto en los resultados del *glycan array*, está presente en algunas bacterias patógenas (Stowell *et al.*, 2010), por lo que podría ser un determinante a tener en cuenta para el desarrollo de terapias antibacterianas. En este sentido, algunos sialilglicoconjugados, como la 3'-sialillactosa, la 6'-sialillactosa o la 6'-sialil(N-acetilactosamina), son clave en el reconocimiento molecular de virus, como el de la gripe. De hecho, diferentes variantes del virus de la gripe reconocen mejor uno u otro derivado sialilado (Gambaryan *et al.*, 1997). Teniendo en cuenta la mayor capacidad de reconocimiento de LSL<sub>150</sub>-His de estructuras de poli-LacNAc con unidades de Neu5Ac unidas mediante enlace  $\alpha(2\rightarrow6)$ , la lectina podría ser de interés en el desarrollo de estrategias terapéuticas y/o de diagnóstico.

Además de una posible aplicación de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> en estrategias de diagnóstico molecular, la especificidad de estas dos lectinas puede ser aprovechada



en el área de la Glicobiología para enriquecer ensayos de tipo *microarray* de lectinas (Hirabayashi *et al.*, 2013) o en proyectos de glicómica celular (Tateno *et al.*, 2010).

Finalmente, con respecto a la importancia de las cadenas de poli-LacNAc en hongos superiores, no hay nada descrito hasta el momento. Por esta razón, extraer conclusiones acerca de la posible función biológica de LSLa y LBL en los respectivos hongos *L. sulphureus* y *L. bicolor* es complicado, teniendo solamente en cuenta el análisis de la especificidad.

#### 7.1.4. Los tréboles $\beta$ fúngicos presentan diferencias en su especificidad

Los ensayos de especificidad llevados a cabo sobre otras lectinas fúngicas con estructura tipo trébol  $\beta$  indican que son específicas de azúcares con unidades de galactosa, una característica común entre prácticamente todos los tréboles  $\beta$  descritos. Sin embargo, muchas de ellas, y a diferencia de lo que ocurre con LSL<sub>150</sub> y LBL<sub>152</sub>, reconocen con mayor afinidad azúcares que presentan N-acetilgalactosamina (GalNAc), en lugar de un simple anillo de galactosa, en posición terminal. Por ejemplo, las lectinas SSA, CNL y RSA reconocen con mucha afinidad el disacárido GalNAc $\beta$ (1 $\rightarrow$ 4)GlcNAc (N,N'-diacetillactosadamina, LacdiNAc o LDN) (Sulzenbacher *et al.*, 2010; Pohleven *et al.*, 2012; Skamnaki *et al.*, 2013), un epítipo poco común en mamíferos, pero que es característico de organismos invertebrados, como helmintos parásitos e insectos (Van den Eijnden *et al.*, 1995; Sasaki *et al.*, 2007). Además, en el caso de SSA, que sólo posee un sitio de unión a carbohidratos, la afinidad disminuye con la aparición de ramificaciones y con el aumento de la longitud de las cadenas de los oligosacáridos, un comportamiento completamente opuesto al descrito en las proteínas de este trabajo. Asimismo, cabe citar que la lectina CCL2, que curiosamente tiene un modo de unión completamente distinto al resto de lectinas, une específicamente residuos de glucosa (Schubert *et al.*, 2012), como el antígeno Le<sup>x</sup>, presente igualmente en algunas especies de gusanos helmintos (Nyame *et al.*, 1998).

El patrón de especificidad observado en LSL<sub>150</sub>-His y LBL<sub>152</sub>-His coincide en gran parte con el descrito para la galectina-3 humana (Seetharaman *et al.*, 1998). Esta galectina se caracteriza por reconocer cadenas de poli-LacNAc y está implicada en la respuesta del sistema inmunológico y en procesos de carcinogénesis (Salomonsson *et al.*, 2010). Además, se ha comprobado que puede reconocer otras estructuras, entre ellas a la del disacárido LDN (van den Berg *et al.*, 2004). Por otro lado, las diferencias que muestra LSL<sub>150</sub>-His en el patrón de reconocimiento de azúcares frente a LBL<sub>152</sub>-His hace que su especificidad sea también parecida a la de otras lectinas. Por ejemplo, la capacidad de unir carbohidratos con

ácido siálico en posición (2→6) es compartida con la citolisina de *Vibrio cholerae* y la he-maglutinina H1N2 de la gripe A, mientras que el reconocimiento del antígeno sanguíneo A es común a la aglutinina WGA del germen de trigo y la lectina fúngica CNL.

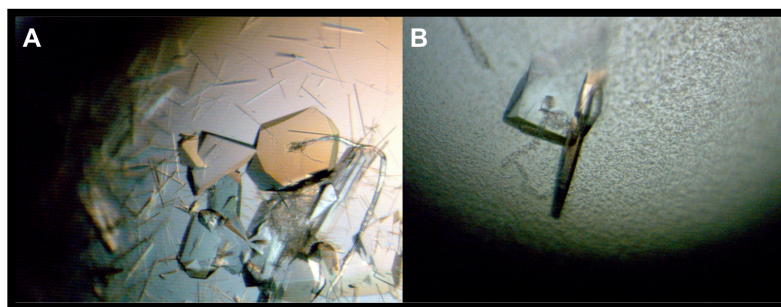
## 7.2. Los complejos con LacNAc revelan nuevas interacciones

Con el propósito de comprender en detalle cuáles son los determinantes de la unión que presentan los dos módulos lectina frente al azúcar LacNAc, unidad fundamental en la construcción de las colas de poli-LacNAc, se llevó a cabo el estudio cristalográfico de los respectivos complejos con dicho disacárido.

### 7.2.1. Cristalización y resolución estructural de los complejos con LacNAc

Los dos módulos lectina, LSL<sub>150</sub> y LBL<sub>152</sub>, se concentraron hasta los 25 mg/mL en tampón TSA y se preincubaron durante 30 minutos con diferentes concentraciones de LacNAc: 20, 40 y 60 mM. Los ensayos de cristalización se llevaron a cabo mediante difusión de vapor en gota colgante a 18 °C, utilizando diferentes relaciones de volumen de proteína:precipitante: 1:1, 1:2 y 2:1. La condición precipitante que se utilizó en cada caso fue la solución en la que se habían obtenido previamente los cristales de LSL<sub>150</sub> y LBL<sub>152</sub> en presencia de lactosa. De esta manera, los mejores cristales del complejo LSL<sub>150</sub>:LacNAc aparecieron en la condición 12% (p/v) PEG 4000, 0,1 M Tris-HCl pH 8,5 y 150 mM acetato de sodio en presencia de 40 mM LacNAc y usando una relación de volúmenes 2:1 (**Figura 6.6A**). Por otra parte, los cristales del complejo LBL<sub>152</sub>:LacNAc de mayor calidad se obtuvieron en la condición 20% (p/v) PEG 6000, 0,1 M acetato de sodio pH 5,0 y 0,2 M cloruro de amonio, utilizando 60 mM LacNAc y una relación de volúmenes 2:1 (**Figura 6.6B**).

Antes de llevar a cabo el experimento de difracción de rayos X, los cristales se sumergieron durante ~10 s en una solución que contenía la condición de cristalización, el ligando LacNAc y un suplemento de un 20% (v/v) de glicerol. Una vez crioprotegidos, los cristales se midieron en la línea ID14-1 del ESRF (Grenoble, Francia), empleando una longitud de onda de 0,939 Å y usando un detector ADSC Q315R. En ambos casos, la recogida de las imágenes se hizo a 100 K y con un ángulo de rotación de 1°.



**Figura 6.6.** Cristales de los complejos de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> con N-acetillactosamina (LacNAc) obtenidos a 18 °C mediante difusión de vapor en gota colgante. **A.** Fotografía bajo luz polarizada de los cristales del complejo LSL<sub>150</sub>:LacNAc aparecidos en 12% (p/v) PEG 4000, 0,1 M Tris-HCl pH 8,5 y 150 mM acetato de sodio en presencia de 40 mM LacNAc. **B.** Cristales del complejo LBL<sub>152</sub>:LacNAc aparecidos en 20% (p/v) PEG 6000, 0,1 M acetato de sodio pH 5,0 y 0,2 M cloruro de amonio en presencia de 60 mM LacNAc.

Los datos cristalográficos fueron procesados con los programas iMosflm (Leslie, 2006) y SCALA (Evans, 2006) del paquete de software CCP4 (Winn *et al.*, 2011). La resolución estructural se llevó a cabo mediante reemplazo molecular con PHASER (McCoy *et al.*, 2007), utilizando las respectivas estructuras de LSL<sub>150</sub> y LBL<sub>152</sub>:Lac como modelos de reemplazo para cada uno de los complejos. Finalmente, la reconstrucción y el refinamiento se hizo con Coot (Emsley y Cowtan, 2004) y PHENIX (Adams *et al.*, 2010), respectivamente. Debido a la buena calidad de los datos, se alcanzaron resoluciones muy elevadas ( $\sim 1$  Å) que permitieron aplicar refinamiento anisótropo al modelo y añadir átomos de hidrógeno a la molécula de proteína, lo que contribuye favorablemente a las estadísticas finales. Las estadísticas del procesamiento de los datos cristalográficos y de la resolución estructural quedan recogidas en la **Tabla 6.3**.

Los cristales de los complejos LSL<sub>150</sub>:LacNAc y LBL<sub>152</sub>:LacNAc presentan los grupos espaciales  $P3_2$  y  $P2_12_12$ , respectivamente. Este resultado, junto con el de las dimensiones de la celda unidad (**Tabla 6.3**), confirma que ambos cristales son equivalentes a los obtenidos en los complejos con lactosa. En ambos casos, hay una molécula de proteína en la unidad asimétrica y el porcentaje de solvente es mayor del 40%, determinado a partir del volumen de Matthews (Kantardjieff y Rupp, 2003). Por último, las estructuras no sufren cambios con respecto a las obtenidas con antelación, lo que significa que el disacárido LacNAc no afecta de ningún modo al plegamiento global de la proteína, ni siquiera al empaquetamiento cristalino.

Parámetros	LSL <sub>150</sub> :LacNAc	LBL <sub>152</sub> :LacNAc
Recogida de datos y procesamiento		
Grupo espacial	$P3_2$	$P2_12_12$
Moléculas en la UA	1	1
Celda unidad (Å)	62,19, 62,19, 37,78	44,54, 52,53, 61,01
Longitud de onda (Å)	0,939	0,939
Resolución (Å)	27–1,14	23–1,06
Reflexiones totales	350.099	134.198
Reflexiones únicas	55.971	60.225
Complejidad	94,9 (88,2)	91,5 (92,2)
Multiplicidad	6,3 (5,9)	2,2 (2,0)
$R_{merge}$ (%)	7,2 (25,9)	6,4 (19,3)
$I/\sigma(I)$	15,8 (6,9)	10,7 (6,4)
Refinamiento		
$R_{work}/R_{free}$	0,136/0,143	0,157/0,175
RMSD de los enlaces (Å)	0,007	0,008
RMSD de los ángulos (°)	1,465	1,544
Factor $B$ medio (Å <sup>2</sup> )	10,41	8,33
Número de átomos		
Proteína	1.325	1.346
Glicerol	6	0
LacNAc	52	78
Agua	289	250
Mapa de Ramachandran		
Favorecidos (%)	98,7	96,3
Permitidos (%)	1,3	3,7
Prohibidos (%)	0	0

\*Los valores de la concha de mayor resolución se muestran en paréntesis

**Tabla 6.3.** Datos de difracción y estadísticas del refinamiento de las estructuras de los complejos de LSL<sub>150</sub> y LBL<sub>152</sub> con N-acetilactosamina (LacNAc).

### 7.2.2. Determinantes estructurales de la unión específica a LacNAc

Al igual que se había observado previamente en los complejos con lactosa, LSL<sub>150</sub> posee sólo dos sitios operativos para unir LacNAc, mientras que LBL<sub>152</sub> sigue teniendo los tres sitios disponibles. El hecho de que se observen moléculas de LacNAc en todos los sitios funcionales valida en parte la idea extraída de los resultados del *glycan array*, en los que se observa mayor unión con los glicanos de tipo poli-LacNAc ramificados, indicando que todos los sitios pueden participar en el reconocimiento dando lugar a una unión multivalente. En ambos casos, los factores  $B$  de los ligandos son semejantes a los de los residuos que están interaccionando directamente con ellos, lo que indica que todos los sitios están completamente ocupados. La excelente calidad de los mapas permite diferenciar, al igual que se hizo durante el estudio de la unión de lactosa, entre las dos formas anoméricas de la N-acetilactosamina ( $\alpha$  y  $\beta$ ). Las ocupaciones de los dos anómeros rondan el 50%, por lo que los sitios de unión aparentemente no muestran predilección por uno u otro anómero.

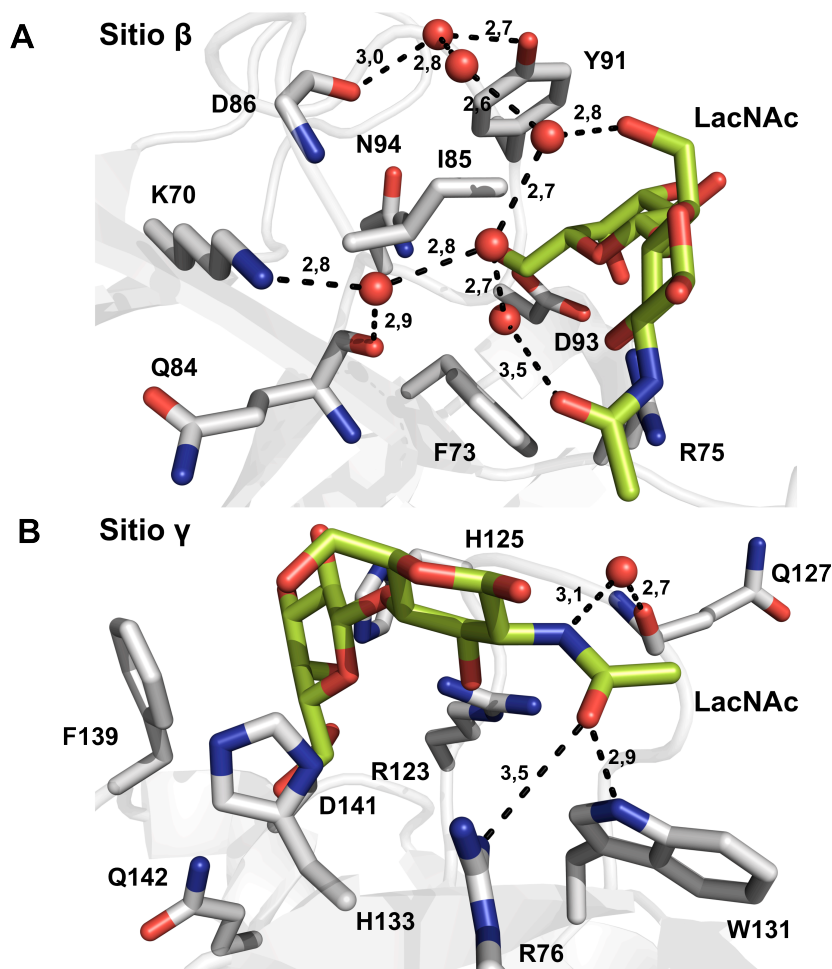
En este sentido, no se observan interacciones específicas que involucren al grupo hidroxilo del carbono anomérico.

En lo que al modo de unión del disacárido LacNAc respecta, se mantienen las mismas interacciones que ya habían sido descritas anteriormente para la lactosa. De hecho, la molécula de LacNAc se orienta de la misma manera que la lactosa en los sitios de unión, estableciendo un apilamiento entre la cara B del anillo de galactosa y un residuo aromático, y formando una red de enlaces de hidrógeno con los residuos del sitio de unión y moléculas de solvente. No obstante, aparecen nuevas interacciones que actúan específicamente sobre el grupo acetamido del anillo de glucosa. La incorporación del grupo acetamido en muchos ligandos (GlcNAc, GalNAc, NeuNAc, etc.) es un elemento clave en la especificidad que muchas lectinas exhiben (Weis y Drickamer, 1996). A diferencia de los grupos hidroxilo, el grupo amida y el oxígeno carbonílico del grupo acetamido tienen una geometría plana y fija. El grupo amida actúa como un dador en el enlace de hidrógeno, mientras que el átomo de oxígeno frecuentemente se comporta como un aceptor del enlace de hidrógeno, interactuando frecuentemente con residuos de serina (Weis y Drickamer, 1996).

#### 7.2.2.1. LSL<sub>150</sub>

En el caso del sitio  $\beta$  del módulo lectina LSL<sub>150</sub> (**Figura 6.7A**), curiosamente, el grupo acetamido no participa en ninguna interacción directamente con residuos del sitio de unión. Éste queda bastante alejado de los aminoácidos más próximos, pero sí que puede participar en una red de moléculas de agua. Esto no implica que estas uniones no sean igual de importantes porque, como ya se vio anteriormente, los enlaces de hidrógeno mediados por moléculas de agua son igual de fuertes que los mediados directamente por residuos del sitio de unión (Toone, 1994). De este modo, el oxígeno carbonílico fija una molécula de agua que a su vez está interconectada con otras moléculas de agua y residuos del sitio de unión, por lo que estas moléculas de agua actuarían como una extensión del propio sitio de unión. En el sitio  $\gamma$  (**Figura 6.7B**), en cambio, el grupo acetamido sí que está involucrado directamente en contactos con residuos del sitio de unión. Por un lado, el nitrógeno amida forma un enlace de hidrógeno con una molécula de agua que está siendo reconocida por el oxígeno carbonílico del enlace peptídico de la Gln-127. Por otro lado, el oxígeno del grupo acetamido forma sendos enlaces de hidrógeno con el N $\epsilon$ -1 del Trp-131 y uno de los N $\eta$  de la Arg-76. Este mayor número de contactos podría repercutir en una mayor afinidad intrínseca del sitio  $\gamma$  por LacNAc, como ya se había visto en el caso de la lactosa. En efecto, estudios previos habían demostrado que solamente el sitio de unión  $\gamma$  se ocupa con moléculas

de LacNAc tras una inmersión rápida (*quick soaking*) de cristales de LSLa en una solución con 20 mM LacNAc. Para observar ocupación también en el sitio  $\beta$  se necesita, sin embargo, una inmersión en una solución más concentrada (100 mM LacNAc) y durante mucho más tiempo (Mancheño *et al.*, 2005).

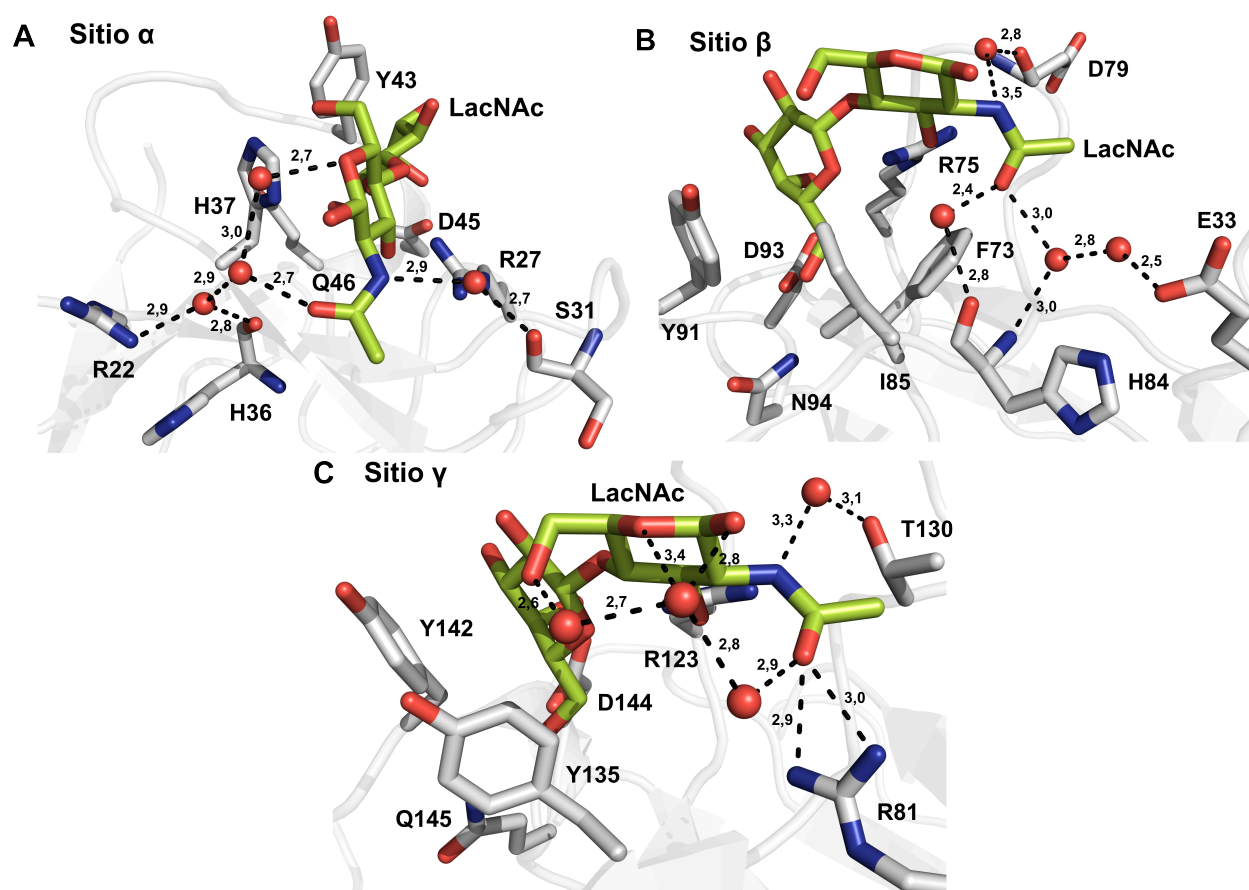


**Figura 6.7.** Vista detallada de la unión de LacNAc a los sitios  $\beta$  (A) y  $\gamma$  (B) de LSL<sub>150</sub>. En ambos casos se muestra el anómero  $\beta$  de la LacNAc (*amarillo*) y los residuos implicados en la unión (*gris*). Además, se representan las moléculas de agua (*esferas rojas*) así como las distancias susceptibles de formar enlaces de hidrógeno (*líneas discontinuas*) que intervengan específicamente en el reconocimiento del grupo acetamido del anillo de glucosa. Solamente se muestran aquellos enlaces que se encontrarían en el intervalo de distancias 2,0–3,5 Å.

#### 7.2.2.2. LBL<sub>152</sub>

El reconocimiento de LacNAc en los sitios de unión de LBL<sub>152</sub> (**Figura 6.8**) también se lleva a cabo mediante la participación de moléculas de agua. De hecho, en los sitios de unión  $\alpha$  y  $\beta$ , no se observan contactos directos del grupo acetamido con ningún aminoácido. En el caso del sitio  $\alpha$  (**Figura 6.8A**), el nitrógeno amida interacciona con la molécula de agua que forma un enlace de hidrógeno con el oxígeno de la cadena principal de la Ser-31 y el oxígeno carbonílico establece un enlace de hidrógeno con una molécula de agua que

forma a su vez más enlaces con otras moléculas del solvente. De igual forma, en el sitio  $\beta$  (**Figura 6.8B**), el nitrógeno amida enlaza con el oxígeno de la cadena principal del Asp-79 mediante un puente con una molécula de agua, y el oxígeno carbonílico establece un puente similar, pero con el residuo de His-84. A diferencia de lo que ocurre en estos dos sitios, en el sitio  $\gamma$  (**Figura 6.8C**), el oxígeno del grupo acetamido forma dos enlaces de hidrógeno directamente con los nitrógeno N $\eta$ -1 y 2 del aminoácido Arg-81, además de reconocer una molécula de agua que está implicada en una red con más moléculas de solvente. Asimismo, el nitrógeno amida interacciona con el residuo de Thr-130 de la misma forma que en los otros dos sitios de unión. Al igual que se ha descrito para el módulo lectina LSL<sub>150</sub>, el mayor número de contactos del sitio de unión  $\gamma$  en LBL<sub>152</sub> podría contribuir a un aumento de afinidad por LacNAc con respecto a los otros dos sitios operativos.



**Figura 6.8.** Vista detallada de la unión de LacNAc a los sitios  $\alpha$  (A),  $\beta$  (B) y  $\gamma$  (C) de LBL<sub>152</sub>. En ambos casos se muestra el anómero  $\beta$  de la LacNAc (*amarillo*) y los residuos implicados en la unión (*gris*). Además, se representan las moléculas de agua (*esferas rojas*) así como las distancias susceptibles de formar enlaces de hidrógeno (*líneas discontinuas*) que intervengan específicamente en el reconocimiento del grupo acetamido del anillo de glucosa. Solamente se muestran aquellos enlaces que se encontrarían en el intervalo de distancias 2,0–3,5 Å.

### 7.3. Caracterización termodinámica de la unión del disacárido LacNAc

Finalmente, la interacción de las dos lectinas con el azúcar LacNAc fue estudiada mediante ITC para evaluar de forma cuantitativa el incremento de afinidad experimentado con respecto a la lactosa, como se desprende de los resultados estructurales y del *glycan array*. Los ensayos se llevaron a cabo a 25 °C en tampón TSA, empleando una concentración final del disacárido LacNAc de 30 mM. La orientación de los picos en el perfil de la unión (**Figuras 6.9 y 6.10**) demuestra en ambos casos que ésta es exotérmica, tal y como ocurría anteriormente con la lactosa. Sin embargo, se observa que con menor concentración de ligando, el sistema alcanza la saturación más rápidamente, lo que sugiere *a priori* que la unión está más favorecida.

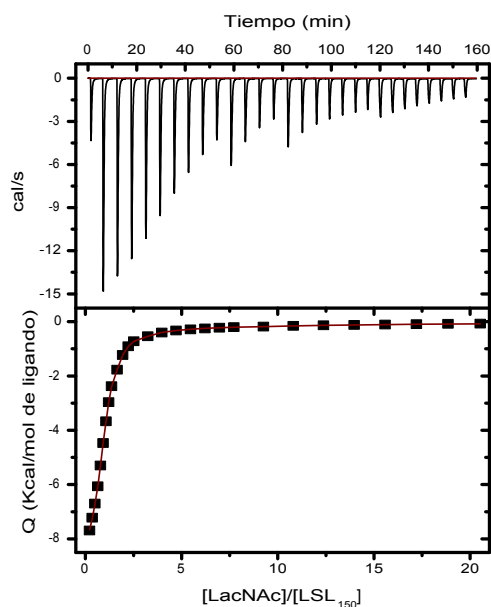
#### 7.3.1. LSL<sub>150</sub>

En el caso particular de LSL<sub>150</sub>, los datos experimentales se ajustan perfectamente a un modelo sencillo de dos sitios de unión independientes. Al igual que se había descrito en el reconocimiento de la lactosa, se distinguen dos sitios diferentes, uno de mayor afinidad y otro de menor. La estequiometría del sitio de mayor afinidad se pudo estimar experimentalmente a un valor de prácticamente 1 ( $0,963 \pm 0,002$ ); los dos sitios de unión identificados se corresponderían con los observados en la estructura cristalográfica del complejo. Como se muestra detalladamente en la Tabla 6.4, el sitio de mayor afinidad presenta una constante de disociación ( $K_{d1}$ ) de  $42,0 \pm 0,7 \mu\text{M}$  ( $\Delta G_1 = -5,97 \pm 0,01 \text{ kcal/mol}$ ), lo que supone prácticamente un aumento de cuatro veces la afinidad observada por la lactosa (**Figura 4.13**). El sitio de menor afinidad, por su parte, experimenta una disminución en la constante de disociación ( $K_{d2}$ ) de aproximadamente la mitad del valor del determinado en la unión de la lactosa, lo que implica, en consecuencia, el doble de afinidad. Al igual que la unión de lactosa, la unión de LacNAc está dirigida entálpicamente, con una contribución entrópica desfavorable. Esto es consistente con el establecimiento de nuevos contactos entre el ligando y los sitios de unión, descritos en el apartado anterior a la luz de la estructura cristalográfica.

Aunque la bondad del ajuste ( $\chi^2$ ) es ligeramente peor, los datos experimentales de la unión de LacNAc por LSL<sub>150</sub> también pueden ajustarse a un modelo secuencial con dos sitios de unión. La aplicación de este modelo, como se desprende de los valores de la Tabla 6.4, muestra unos parámetros termodinámicos muy parecidos a los obtenidos con el otro modelo de ajuste. La diferencia más significativa se encuentra en los valores determinados



para el sitio de menor afinidad, ya que en este caso la unión de LacNac sería equivalente a la de la lactosa en términos de afinidad. Esta discrepancia con el estudio cristalográfico valida la utilización del modelo de dos sitios de unión independientes para analizar los datos de ITC.



**Figura 6.9.** Análisis mediante ITC de la unión en solución de LacNac a LSL<sub>150</sub>. Se muestra la curva de valoración, así como el ajuste de los datos (rojo) con el modelo de dos sitios de unión independientes.

#### LSL<sub>150</sub>:LacNac

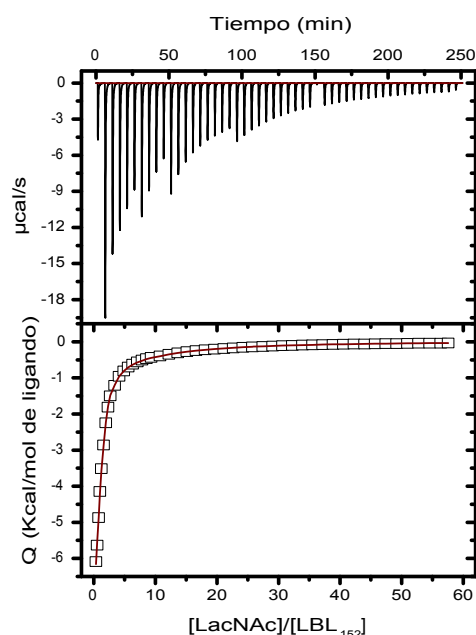
	<i>Dos sitios de unión independientes</i>		<i>Dos sitios de unión secuenciales</i>	
	Alta afinidad	Baja afinidad	Alta afinidad	Baja afinidad
<i>n</i>	0,963 ± 0,002	1	1	1
<i>K<sub>d</sub></i>	42,0 ± 0,7 μM	6,2 ± 0,6 mM	43 ± 2 μM	19 ± 8 mM
$\Delta G$ (kcal/mol)	- 5,97 ± 0,01	- 3,01 ± 0,06	- 5,96 ± 0,03	- 2,4 ± 0,2
$\Delta H$ (kcal/mol)	- 9,18 ± 0,03	- 7,0 ± 0,4	- 9,00 ± 0,08	- 14 ± 5
$T\Delta S$ (kcal/mol)	- 3,21 ± 0,04	- 4,0 ± 0,5	- 3,0 ± 0,1	- 12 ± 5

**Tabla 6.4.** Análisis termodinámico de la unión a LacNac mediante ITC de LSL<sub>150</sub>. Los diferentes parámetros termodinámicos (la estequiometría (*n*), la constante de disociación (*K<sub>d</sub>*), la energía de Gibbs ( $\Delta G$ ), la entalpía de unión ( $\Delta H$ ) y el cambio de entropía ( $T\Delta S$ )) se han determinado aplicando dos modelos de ajuste distintos: dos sitios de unión independientes y dos sitios de unión secuenciales.

Independientemente del modelo usado, el sitio  $\gamma$  se correspondería con mayor probabilidad con el sitio de unión de mayor afinidad, determinando la especificidad de la lectina. Este hecho nos permite lanzar la hipótesis de que el sitio  $\gamma$  probablemente establece el primer contacto con los glicanos de tipo poli-LacNac, favoreciéndose así el establecimiento de interacciones entre el sitio  $\beta$  y el resto de la cadena polisacáridica y traduciéndose en definitiva en un aumento sustancial de la afinidad final.

7.3.2. LBL<sub>152</sub>

Los datos de ITC obtenidos para LBL<sub>152</sub> se ajustan también a un modelo de dos tipos de sitios de unión independientes. En este caso y a diferencia de lo que ocurría con la lactosa, la estequiometría se estima experimentalmente a 1 ( $1,11 \pm 0,01$ ) para el sitio de mayor afinidad. Por esta razón y en base a la estructura cristalográfica, la estequiometría se fijó en 2 para los sitios de menor afinidad. El análisis de los parámetros termodinámicos aplicando este modelo (Tabla 6.5) revela una afinidad elevada por la LacNAc para el sitio de mayor afinidad [ $K_{d1} = 76 \pm 5 \mu\text{M}$  ( $\Delta G_1 = -5,62 \pm 0,04 \text{ kcal/mol}$ )] y del mismo orden de magnitud que la obtenida para el sitio  $\gamma$  de LSL<sub>150</sub>. Si se compara el valor de esta constante con el obtenido para la unión de lactosa en los sitios de mayor afinidad (**Figura 4.41**), el ligando LacNAc presenta una afinidad 20 veces superior que la lactosa. En relación con los sitios de menor afinidad, estos también unen mejor LacNAc que lactosa; de hecho, el valor de la constante de disociación [ $K_{d2} = 3,5 \pm 0,3 \text{ mM}$  ( $\Delta G_1 = -3,35 \pm 0,05 \text{ kcal/mol}$ )] es aproximadamente un tercio del obtenido en la unión de la lactosa. En ambos tipos de sitios, la interacción con el disacárido transcurre con una pérdida de entropía ( $T\Delta S < 0$ ), pero con una contribución entálpica favorable.



**Figura 6.10.** Análisis mediante ITC de la unión en solución de LacNAc a LBL<sub>152</sub>. Se muestra la curva de valoración, así como el ajuste de los datos (rojo) con el modelo de dos sitios de unión independientes.

Si se aplica en este caso un modelo de tipo secuencial, el mejor ajuste se consigue cuando se definen tres sitios de unión independientes. En este caso, el sitio de mayor afinidad presenta una constante aún más baja [ $K_{d1} = 32 \pm 2 \mu\text{M}$  ( $\Delta G_1 = -6,12 \pm 0,04 \text{ kcal/mol}$ )] y el sitio de menor afinidad posee una constante parecida a la obtenida con el otro modelo

[ $K_{d3} = 5,0 \pm 0,5$  mM ( $\Delta G_1 = -3,13 \pm 0,06$  kcal/mol)]. Es interesante el sitio de afinidad media, puesto que exhibe mayor afinidad que la observada por la lactosa, pero considerablemente menor que la del sitio de mayor afinidad [ $K_{d2} = 360 \pm 40$   $\mu$ M ( $\Delta G_1 = -4,69 \pm 0,06$  kcal/mol)].

LBL <sub>152</sub> :LacNAc			
Dos sitios de unión independientes			
	Alta afinidad	Baja afinidad	
$n$	1,11 $\pm$ 0,01	2	
$K_d$	76 $\pm$ 5 $\mu$ M	3,5 $\pm$ 0,3 mM	
$\Delta G$ (kcal/mol)	-5,62 $\pm$ 0,04	-3,35 $\pm$ 0,05	
$\Delta H$ (kcal/mol)	-8,6 $\pm$ 0,2	-7,5 $\pm$ 0,2	
$T\Delta S$ (kcal/mol)	-3,0 $\pm$ 0,2	-4,2 $\pm$ 0,3	

LBL <sub>152</sub> :LacNAc			
Tres sitios de unión secuenciales			
	Alta afinidad	Afinidad media	Baja afinidad
$n$	1	1	1
$K_d$	32 $\pm$ 2 $\mu$ M	360 $\pm$ 40 $\mu$ M	5,0 $\pm$ 0,5 mM
$\Delta G$ (kcal/mol)	-6,12 $\pm$ 0,04	-4,69 $\pm$ 0,06	-3,13 $\pm$ 0,06
$\Delta H$ (kcal/mol)	-7,45 $\pm$ 0,09	-4,4 $\pm$ 0,3	-13,7 $\pm$ 0,5
$T\Delta S$ (kcal/mol)	-1,3 $\pm$ 0,1	0,3 $\pm$ 0,4	-10,5 $\pm$ 0,6

**Tabla 6.5.** Análisis termodinámico de la unión a LacNAc mediante ITC de LBL<sub>152</sub>. Los diferentes parámetros termodinámicos [la estequiometría ( $n$ ), la constante de disociación ( $K_d$ ), la energía de Gibbs ( $\Delta G$ ), la entalpía de unión ( $\Delta H$ ) y el cambio de entropía ( $T\Delta S$ )] se han determinado aplicando dos modelos de ajuste distintos: dos sitios de unión independientes y tres sitios de unión secuenciales.

Teniendo en cuenta las nuevas interacciones que aparecen cuando el ligando es LacNAc, es difícil identificar cuál sería el sitio de mayor afinidad. Por un lado, éste podría seguir siendo el sitio del subdominio  $\alpha$  porque el número de contactos que se establecen con el ligando es más alto, como ya se había observado. Sin embargo, también podría corresponderse con el sitio de unión  $\gamma$  porque es el único que establece contactos directos entre el grupo acetamido y los residuos del sitio de unión. Desde luego, el sitio  $\beta$  sería probablemente un sitio de baja afinidad, ya que no sólo tiene menor número de contactos con la proteína, sino que además, como se explicó en el apartado 2.7 de *Resultados y Discusión*, las dos cadenas hidrofóbicas de los aminoácidos Phe-73 e Ile-85 podrían interferir negativamente en el reconocimiento. En este contexto, tanto el sitio de unión  $\alpha$  como el  $\gamma$  podrían entonces corresponderse con el sitio de alta afinidad. La evaluación de la energética de los enlaces de hidrógeno que se establecen en ambos sitios de unión podría arrojar luz en este

asunto, pero llevar a cabo este análisis es complicado, teniendo en cuenta que las energías asociadas a cada enlace dependen fuertemente del entorno en el que se encuentren (Fersht, 1987). Generalmente, el número de interacciones va ligado con una mayor afinidad. Por ejemplo, las lectinas CCL2 y PSL presentan numerosos enlaces de hidrógeno (con las cadenas laterales de los residuos del sitio de unión y con grupos de la cadena polipeptídica) e interacciones hidrofóbicas, lo que se traduce en constantes de afinidad por sus ligandos próximas a 1  $\mu\text{M}$  (Tateno *et al.*, 2004; Schubert *et al.*, 2012). En este sentido, el sitio de unión  $\alpha$  podría considerarse el sitio de mayor afinidad, al igual que se dedujo en el estudio de la unión de lactosa.

Aunque las constantes de disociación determinadas para la unión de LacNAc reflejan un aumento considerable de la afinidad con respecto a lo observado para la lactosa, el rango de afinidades sigue estando por debajo del descrito para otras lectinas con sus ligandos específicos. Este comportamiento es típico de lectinas que compensan la baja afinidad con la multivalencia con el fin de aumentar la avidez por la unión de ligandos multivalentes (Dam y Brewer, 2010). En efecto, éste posiblemente sea el caso de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub>, ya que pese a tener unas afinidades relativamente bajas, unen muy bien ligandos multivalentes con largas cadenas de poli-LacNAc, como se deduce de los resultados del *array*. La baja afinidad también está relacionada con la promiscuidad, es decir, generalmente las lectinas que tienen constantes de disociación más bajas suelen presentar una especificidad muy exquisita (Imberthy *et al.*, 2005). Por lo tanto, los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> previsiblemente deberían tener mayor capacidad que otras lectinas para acomodar glicanos y glicoconjugados con mayor variedad estructural. A la vista del *glycan array*, por ejemplo, se observa que LSL<sub>150</sub> une relativamente bien oligosacáridos que están sialilados, fucosilados o que poseen residuos de GlcNAc en posición terminal.

## **8. LSL<sub>150</sub> y LBL<sub>152</sub> podrían reconocer epítomos de insectos y nemátodos**

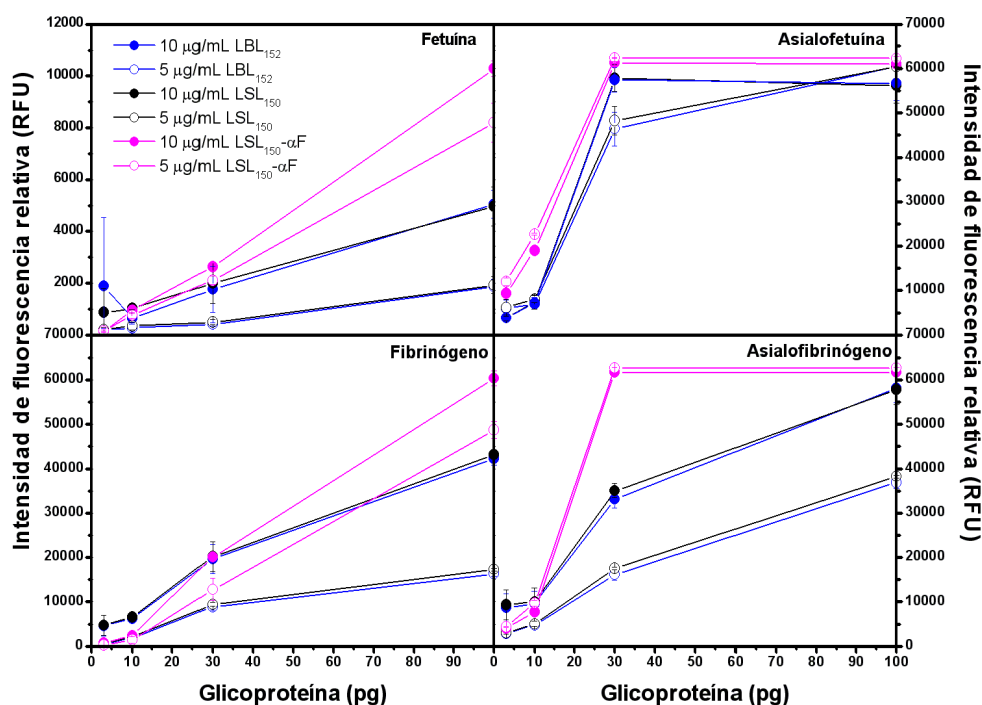
Es sabido que los cuerpos fructíferos de los hongos superiores contienen un gran número de proteínas tóxicas involucradas en los mecanismos de defensa frente a parásitos y depredadores. De hecho, muchas lectinas de cuerpos fructíferos poseen actividad insecticida y nematotóxica. Por ejemplo, las lectinas con plegamiento tipo trébol  $\beta$  procedentes de los hongos *S. sclerotiorum* (SSA) y *R. solani* (RSA) son tóxicas frente al pulgón *Acyrtosiphon pisum* (Hamshou *et al.*, 2010) y el gusano conocido como Rosquilla Negra (*Spodoptera*

*littoralis*) (Hamshou *et al.*, 2013), respectivamente. En ambos casos se demostró que las lectinas se unen al epitelio del intestino medio. También se ha demostrado que esta toxicidad puede ser dependiente de la multivalencia. Así, por ejemplo, la lectina CNL sólo es tóxica frente a *Caenorhabditis elegans* cuando se encuentra en forma dimérica (Pohleven *et al.*, 2012). En el caso de la lectina de *M. oreades* (MOA), su toxicidad es dependiente de dos factores: el reconocimiento de la proteína diana por el módulo lectina N-terminal y la actividad enzimática del dominio C-terminal (Wohlschlager *et al.*, 2011). La identificación del glicoepítipo que reconocen estas lectinas no es tarea sencilla, puesto que hoy en día no se conoce en detalle el glicoma de organismos invertebrados. Recientemente, se ha descubierto, por ejemplo, que la lectina CCL2 de *C. cinerea* une específicamente N-glicanos fucosilados en posición  $\alpha(1\rightarrow3)$  típicos de nemátodos como *C. elegans* (Schubert *et al.*, 2012).

Las lectinas que se expresan en los cuerpos fructíferos y cuya función está estrechamente relacionada con los mecanismos de defensa de los hongos se conocen como lectinas de los cuerpos fructíferos o *fruiting body lectins* (FBL). Teniendo en cuenta que LSLa y LBL se expresan mayoritariamente en los cuerpos fructíferos de los hongos *L. sulphureus* y *L. bicolor*, respectivamente, podrían ser nuevos miembros de este grupo de lectinas. En el caso de LSLa ya se ha demostrado que posee actividad aglutinante y hemolítica (Tateno y Goldstein, 2003), por lo que su papel como mecanismo de defensa es más que probable. En el caso de LBL, la determinación de su función biológica es más complicada, puesto que sólo se produce recombinantemente el módulo N-terminal y la secuencia de la región C-terminal no muestra similitud con ninguna otra proteína. No obstante, el análisis cristalográfico de los complejos con lactosa y LacNAc, así como los resultados del *glycan array*, demuestran que ambas lectinas tienen una especificidad clara por cadenas de poli-LacNAc, pero podrían acomodar otros oligosacáridos con facilidad. Por ejemplo, el azúcar LDN (N, N'-diacetillactosadiamina), un glicoepítipo común en organismos invertebrados como el gusano *Schistosoma mansoni* (Nyame *et al.*, 2002), tiene una estructura química idéntica a la del disacárido LacNAc. Su única diferencia radica en la presencia de un grupo acetamido que sustituye al hidroxilo en posición 2 del anillo de galactosa. Dado que ni LSLa ni LBL tienen residuos en los sitios de unión que reconozcan específicamente al oxígeno O2 de la galactosa, el reconocimiento de este azúcar por parte de las lectinas podría verse favorecido.

## 8.1. Los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> son capaces de unir el glicopéptido LDN

A pesar de que se intentaron obtener las estructuras cristalográficas de los complejos de ambas lectinas con el disacárido LDN, fue imposible obtener cristales tanto por cocrystalización como por inmersión o *soaking*, ya que los cristales se deterioraban rápidamente cuando se sumergían en la solución de cristalización con el azúcar en cuestión. Por esta misma razón, se realizó un *microarray* para determinar mediante ensayos de competición si las lectinas tienen la capacidad de unir este azúcar. Como estos ensayos permiten testar simultáneamente más de un ligando, también se probaron el azúcar 3'-sialillactosa (3'-SL) y el trisacárido Le<sup>x</sup>, un antígeno sanguíneo que aparece asociado a glicoproteínas en algunos organismos invertebrados (Nyame *et al.*, 1998).

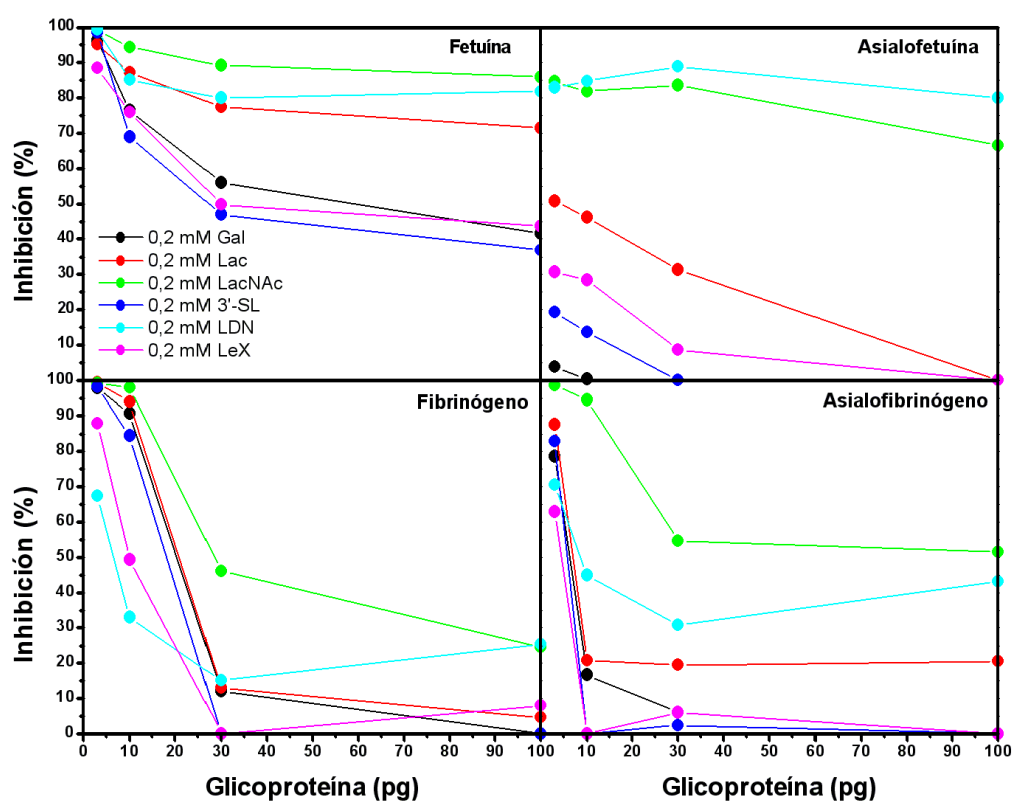


**Figura 6.11.** Perfiles de unión de las lectinas biotiniladas LSL<sub>150</sub> (negro), LBL<sub>152</sub> (azul) y LSL<sub>150</sub>-αF (rosa) a diferentes glicoproteínas obtenidos a partir de medidas de fluorescencia relativa (RFU). Los experimentos se llevaron a cabo por triplicado y utilizando dos concentraciones diferentes de lectina (5 y 10 μg/mL).

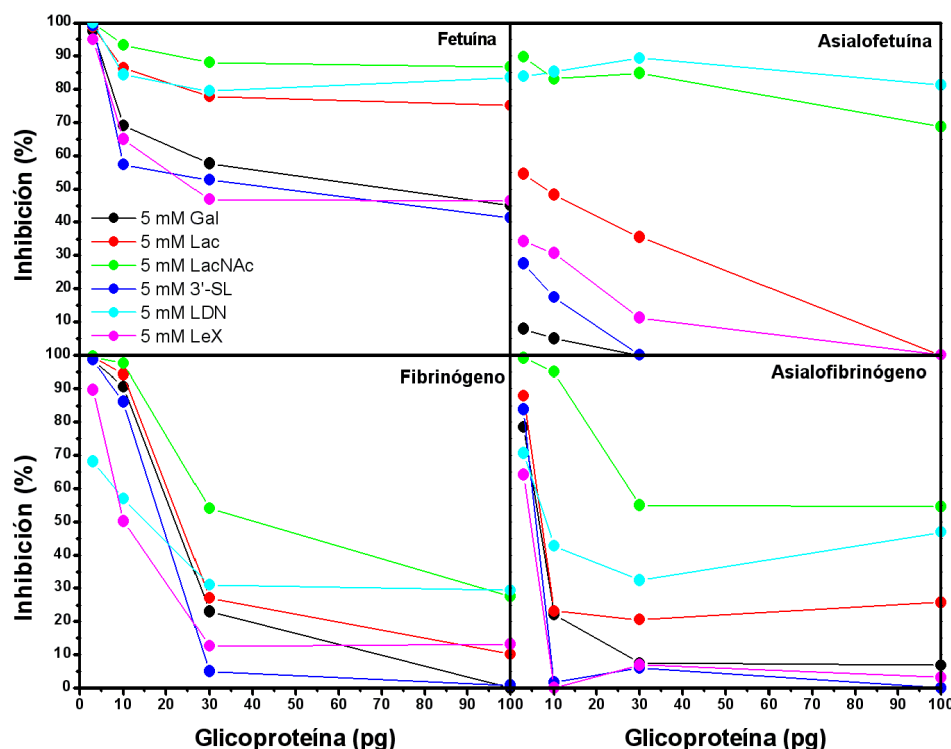
En primer lugar, antes de llevar a cabo los ensayos de competición, se determinó la capacidad de unión de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> a cuatro glicoproteínas modelo: fetuina, asialofetuina, fibrinógeno y asialofibrinógeno (**Figura 6.11**). La unión a estas glicoproteínas permite evaluar el porcentaje de inhibición asociado a los diferentes ligandos. Como se observa en la **Figura 6.11**, las dos lectinas presentan un modo de unión bastante similar. Reconocen muy bien al fibrinógeno y al asialofibrinógeno, lo que es consistente con

los resultados vistos hasta ahora porque estas dos glicoproteínas poseen cadenas de poli-LacNAc (Townsend *et al.*, 1982). En cambio, la unión a fetuína es menor. En este caso, esta glicoproteína posee una mezcla de carbohidratos, entre ellos ácido siálico, galactosa y hexosaminas, principalmente (Spiro, 1960). Cuando los residuos de ácido siálico de esta proteína son eliminados (asialofetuína), la unión mejora considerablemente, siendo similar a la observada para el fibrinógeno y el asialofibrinógeno. Posiblemente, en la asialofetuína, los residuos de galactosa quedan más expuestos, favoreciéndose así el reconocimiento por parte de las lectinas. Como es de esperar, existe una relación directa entre la unión a las glicoproteínas y la concentración de las lectinas.

En el ensayo de unión a las glicoproteínas modelo, también se estudió el triple mutante LSL<sub>150</sub>- $\alpha$ F (Ser29His, Ser37His, Ala43Phe). Al igual que ocurre con la proteína silvestre, se observa que LSL<sub>150</sub>- $\alpha$ F se une muy bien a todas las glicoproteínas con la excepción de la fetuína, de hecho, con más intensidad que la proteína silvestre, puesto que a menores concentraciones de proteína (5  $\mu$ g/mL) la intensidad medida es mayor. Este hecho podría deberse a la multivalencia, puesto que la incorporación de un sitio extra de unión (sitio  $\alpha$ ), aumenta en consecuencia la avidéz que posee por las cadenas de glicanos de las glicoproteínas.



**Figura 6.12.** Perfiles de inhibición de la unión de la lectina LSL<sub>150</sub> biotinilada (10  $\mu$ g/mL) a diferentes glicoproteínas obtenidos a partir de medidas de fluorescencia relativa (RFU). La inhibición se llevó a cabo en presencia de diferentes azúcares a una concentración de 0,2 mM. Todos los ensayos se realizaron por triplicado.



**Figura 6.13.** Perfiles de inhibición de la unión de la lectina LBL<sub>152</sub> biotinilada (10  $\mu\text{g}/\text{mL}$ ) a diferentes glicoproteínas obtenidos a partir de medidas de fluorescencia relativa (RFU). La inhibición se llevó a cabo en presencia de diferentes azúcares a una concentración de 5 mM. Todos los ensayos se realizaron por triplicado.

Una vez se hubo comprobado que las glicoproteínas modelo son reconocidas por las lectinas, se llevaron a cabo los ensayos de competición con azúcares en solución. La concentración de los azúcares se escogió en función de las constantes de afinidad a lactosa determinadas por ITC. De esta manera, las concentraciones utilizadas fueron de 0,2 y 5 mM para LSL<sub>150</sub> y LBL<sub>152</sub>, respectivamente, lo que equivale aproximadamente al doble de la constante de disociación del sitio de mayor afinidad en cada caso. A la vista de los resultados de los ensayos de competición (**Figuras 6.12 y 6.13**), se corrobora de nuevo que el disacárido LacNAc es el que mejor se une, puesto que ejerce una mayor inhibición de la unión a las cuatro glicoproteínas. Además, como cabría esperar, el siguiente azúcar que exhibe un mayor porcentaje de inhibición asociado es LDN. En el caso de la unión a fetúina y asialofetúina, el porcentaje de inhibición por parte de LacNAc y LDN se mantiene elevado incluso en presencia de altas cantidades de glicoproteína inmovilizada (se observan inhibiciones superiores al 70% en presencia de hasta 100 pg de glicoproteína). Sin embargo, en el fibrinógeno y en el asialofibrinógeno, el comportamiento es diferente ya que la inhibición decae rápidamente. Este resultado podría deberse a la elevada especificidad que poseen las lectinas para unirse a las cadenas de poli-LacNAc. Hay que considerar que en este tipo de ensayos el escenario es complejo, las lectinas alcanzan un equilibrio en el que con-



tribuyen tanto las constantes de asociación de los azúcares en solución como las constantes de unión a los glicoproteínas inmovilizadas. Además, y debido a la multivalencia, la avidéz por cadenas de glicanos próximas entre sí es un factor importante que debe tenerse en cuenta a la hora de comparar la unión entre las glicoproteínas y los azúcares en solución.

En relación con el resto de azúcares empleados en el ensayo, solamente la lactosa presenta porcentajes de inhibición relativamente altos en la unión a fetuína (superiores al 70%). Por otro lado, aunque el antígeno Le<sup>x</sup> parece inhibir la unión a fetuína y fibrinógeno, apenas se observa inhibición en las glicoproteínas desializadas, que son las que más afinidad muestran por las lectinas (menor de un 30% de inhibición en presencia de 100 pg de glicoproteína inmovilizada).

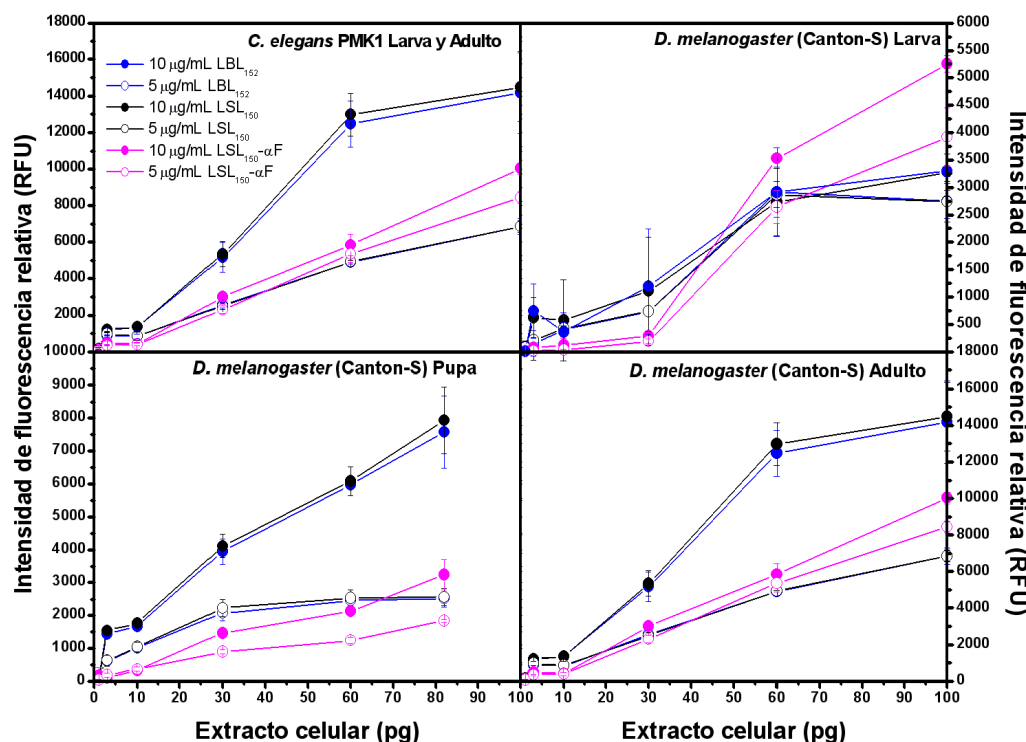
En conclusión, los ensayos de competición han demostrado que los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> tienen la capacidad de unir LDN en solución lo cual apoya su posible rol en los mecanismos de defensa de los hongos.

## 8.2. Extractos celulares de *C. elegans* y *D. melanogaster* son reconocidos por los módulos lectina

Con la intención de avanzar en el conocimiento de la función biológica de ambas lectinas, se llevó a cabo un estudio para verificar si LSL<sub>150</sub> y LBL<sub>152</sub> tienen la capacidad de unirse a extractos celulares de dos organismos modelo, el insecto *D. melanogaster* (Canton-S) y el nemátodo *C. elegans* (PMK1). En el caso de *D. melanogaster*, se trabajó con tres extractos celulares procedentes de tres estadios distintos de maduración del insecto: larva, pupa y adulto. El extracto de *C. elegans*, en cambio, estaba formado por una mezcla de ejemplares del estadio larva y adulto. Puesto que el glicoma varía a lo largo del ciclo vital, es importante trabajar con los diferentes estadios para averiguar en cual de ellos se podría expresar el glicoeptótopo de interés.

En la **Figura 6.14** se muestran los perfiles de unión a los diferentes extractos celulares. En general, se observa que la unión a los extractos celulares es proporcional a la cantidad de extracto inmovilizado en la membrana de nitrocelulosa, tendencia observada igualmente en la unión a las glicoproteínas modelo. Por otro lado, es destacable que los valores de intensidad de fluorescencia más altos se alcanzan únicamente en los extractos de *C. elegans* y en el del estadio adulto de *D. melanogaster*, lo que podría interpretarse como un mecanismo de defensa frente a organismos adultos. Además, por otro lado, es interesante ver cómo ambas lectinas muestran un patrón de especificidad similar. Sorpren-

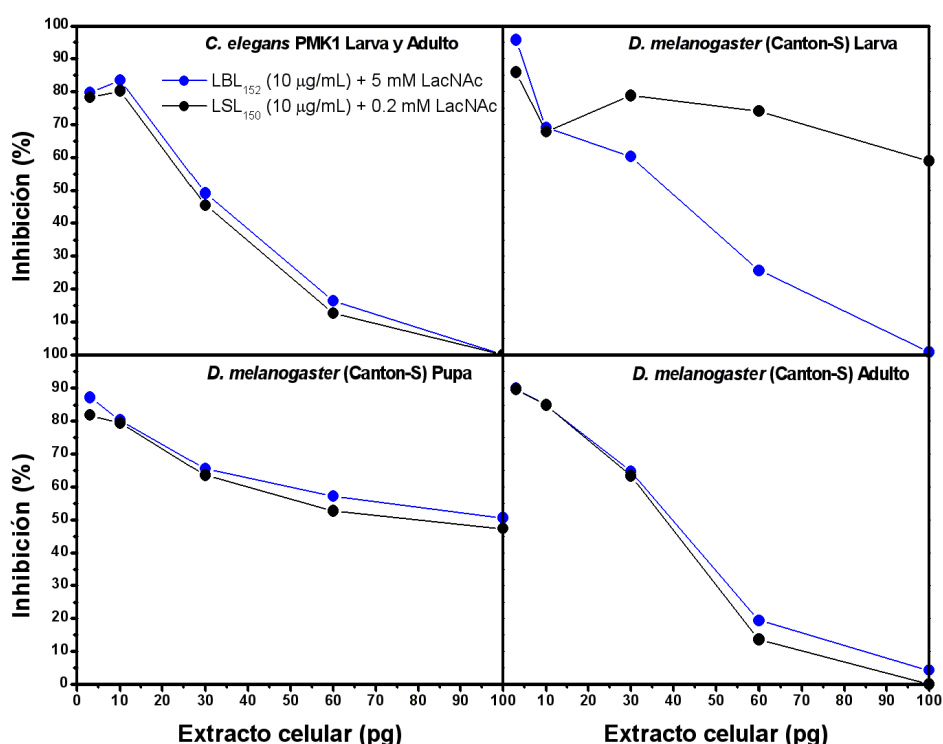
dentemente, el triple mutante de LSL<sub>150</sub> (LSL<sub>150</sub>-αF), aunque unía considerablemente mejor las glicoproteínas modelo, tiene menor capacidad para reconocer los extractos celulares que la proteína silvestre (los valores de intensidad de fluorescencia son más bajos a concentraciones más altas).



**Figura 6.14.** Perfiles de unión de las lectinas biotiniladas LSL<sub>150</sub> (negro), LBL<sub>152</sub> (azul) y LSL<sub>150</sub>-αF (rosa) a diferentes extractos celulares de *Drosophila melanogaster* y *Caenorhabditis elegans* obtenidos a partir de medidas de fluorescencia relativa (RFU). Los experimentos se llevaron a cabo por triplicado y utilizando dos concentraciones diferentes de lectina (5 y 10 μg/mL).

Teniendo en cuenta que los extractos celulares están constituidos por una mezcla heterogénea de proteínas, lípidos, glicoconjugados, etc., se realizaron ensayos de competición en presencia del azúcar LacNAc para comprobar que la unión a los extractos es dependiente principalmente de interacciones lectina-azúcar y no de otro tipo de interacciones, como las de tipo proteína-proteína. En la **Figura 6.15** se representan los perfiles de inhibición, donde se muestra nuevamente que la inhibición es dependiente de la cantidad de extracto celular inmovilizado en la membrana. En los extractos celulares donde se había visto previamente mayor unión (*C. elegans* y *D. melanogaster* adulto), la inhibición decae rápidamente hasta valores inferiores al 10%. Teniendo en cuenta que en las glicoproteínas modelo se veía una inhibición considerable incluso en presencia de 100 pg de muestra, el resultado con los extractos podría indicar que existen interacciones entre las lectinas y los extractos celulares que probablemente no estén mediadas por interacciones de tipo lecti-

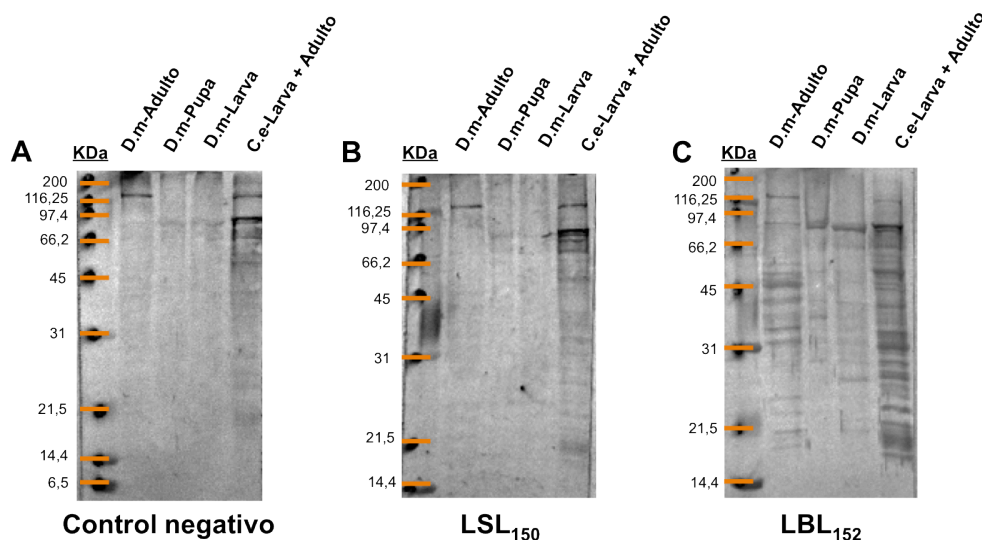
na-azúcar, aunque este comportamiento también podría estar justificado por la presencia de un glicoepítipo muy específico en los extractos celulares, cuya afinidad fuese superior a la determinada por LacNAc en solución. En el caso de *D. melanogaster*, concretamente en los estadios de larva y pupa, sí que se observa inhibición (> 50%) en presencia del disacárido LacNAc, incluso con mucha cantidad de extracto celular, lo que indicaría un reconocimiento mediado básicamente por interacciones lectina-azúcar.



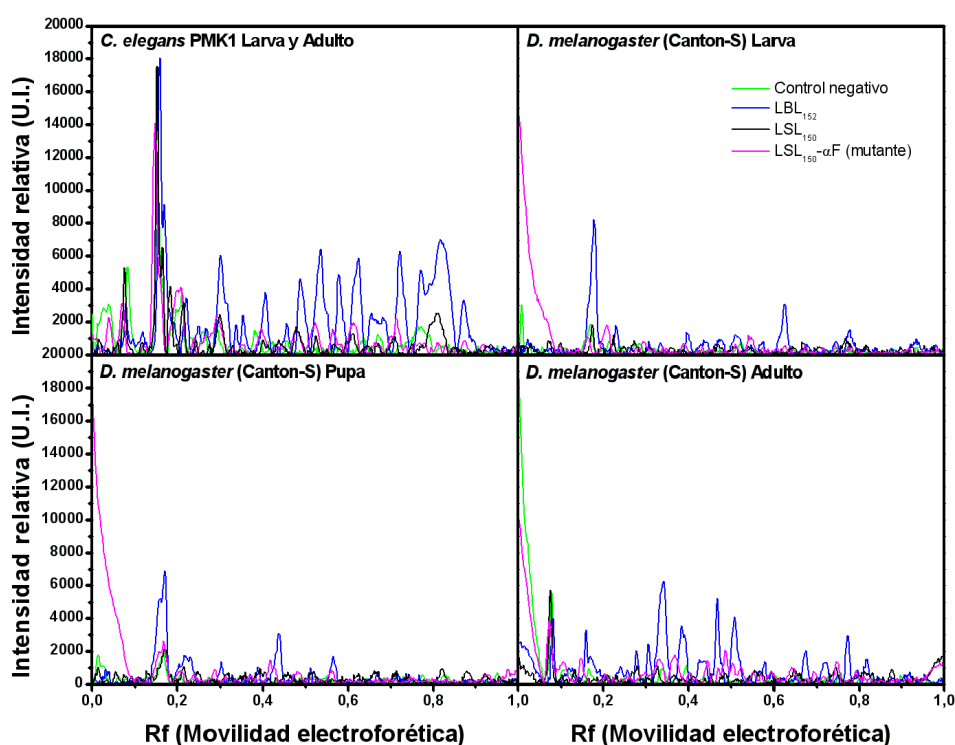
**Figura 6.15.** Perfiles de inhibición de la unión de las lectina biotiniladas LSL<sub>150</sub> (negro) y LBL<sub>152</sub> (azul) a diferentes extractos celulares de *Drosophila melanogaster* y *Caenorhabditis elegans* obtenidos a partir de medidas de fluorescencia relativa (RFU). Los ensayos de inhibición se realizaron por triplicado, utilizando las lectinas a una concentración final de 10 µg/mL y en presencia de N-acetilactosamina (LacNAc) a 0,2 mM para LSL<sub>150</sub> y 5 mM para LBL<sub>152</sub>.

Con la intención de comprender mejor la interacción de las lectinas con los extractos celulares de *D. melanogaster* y *C. elegans*, se llevó a cabo un *lectin blot* sobre ellos (apartado 3.6 de *Materiales y Métodos*) después de haberse realizado una electroforesis monodimensional de tipo SDS-PAGE (**Figura 6.16**). Los resultados de este experimento evidencian algo que no quedaba del todo claro en el *microarray* y es que LSL<sub>150</sub> y LBL<sub>152</sub> tienen un patrón de reconocimiento de los extractos completamente diferente. Mientras que LSL<sub>150</sub> y el control negativo con estreptavidina son equivalentes a la vista de los resultados (**Figuras 6.16A y B**), LBL<sub>152</sub> se une a un mayor número de proteínas (**Figura 6.16C**). Asimismo, el densitometrado de las bandas (**Figura 6.17**) demostró efectivamente que LBL<sub>152</sub> tiene un patrón específico en el que hay proteínas que son reconocidas única y exclusivamente por

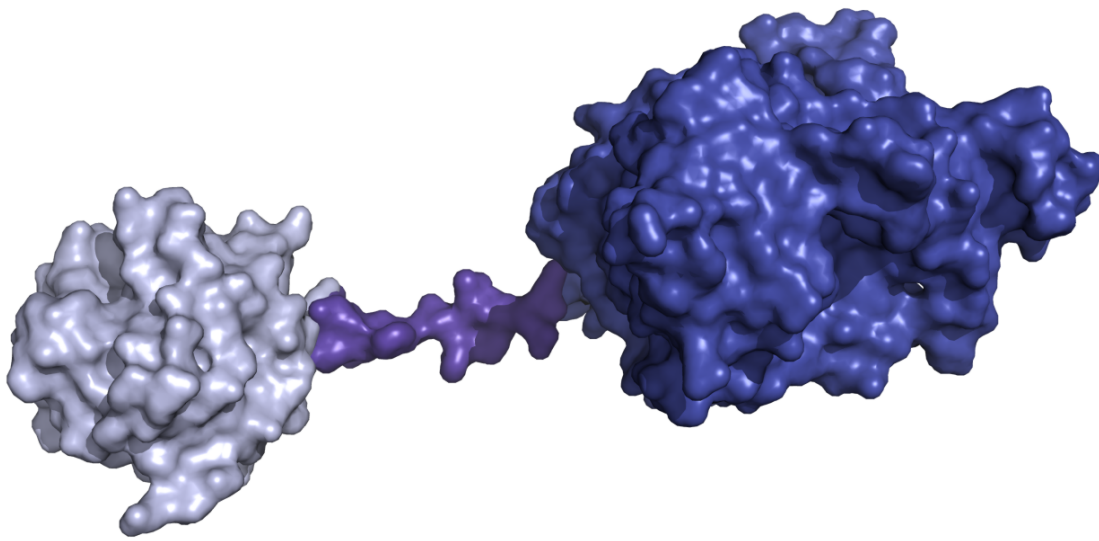
la lectina de *L. bicolor*. Este resultado, por lo tanto, daría algunas pistas más acerca de la función que LBL podría desempeñar en la naturaleza como una toxina vinculada a los mecanismos de defensa del hongo. Sin lugar a dudas, el aislamiento de la proteína completa a partir de los cuerpos fructíferos del hongo así como estudios *in vivo* son necesarios para ahondar más en estos resultados y corroborar definitivamente esta hipótesis (Künzler *et al.*, 2010).



**Figura 6.16.** Patrones de reconocimiento de los extractos celulares de *Drosophila melanogaster* (D.m.) y *Caenorhabditis elegans* (C.e.) obtenidos mediante lectin blotting. En A se muestra el control negativo, resultado de la unión inespecífica de la estreptavidina, mientras que en B y en C se muestra el resultado del reconocimiento llevado a cabo por parte de las lectinas biotiniladas LSL<sub>150</sub> y LBL<sub>152</sub>, respectivamente.



**Figura 6.17.** Análisis por densitometraje del reconocimiento de los extractos celulares de *Drosophila melanogaster* y *Caenorhabditis elegans* por parte de las lectinas biotiniladas LSL<sub>150</sub> y LBL<sub>152</sub>. Asimismo, se ha incluido el perfil de unión del mutante triple LSL<sub>150</sub>-αF.



## **RESULTADOS Y DISCUSIÓN**

**Aplicaciones biotecnológicas  
novedosas**





## RESULTADOS Y DISCUSIÓN

### Aplicaciones biotecnológicas novedosas

#### 9. Producción de proteínas de fusión con dominios trébol $\beta$

Una de las principales características que se han descrito en este trabajo sobre el módulo lectina N-terminal de LSLa (LSL<sub>150</sub>) es que actúa como una unidad de plegamiento independiente. Además, se ha comprobado que este módulo lectina es capaz de solubilizar al módulo C-terminal formador de poros (PFM) y mantener soluble, de esta manera, la proteína completa en una solución acuosa, a lo cual indudablemente influye también en el estado oligomérico de la toxina (Mancheño *et al.*, 2005). A este comportamiento hay que añadir el hecho de que la expresión del módulo lectina en *E. coli* va acompañada de un rendimiento de producción muy elevado, lo que indica que LSL<sub>150</sub> tiene una secuencia de nucleótidos que permite una alta tasa de traducción y que el ambiente de *E. coli* es adecuado para el correcto plegamiento de la lectina. Todas estas propiedades son las que, de hecho, debe cumplir una etiqueta de fusión para actuar como un agente solubilizante (Malhotra, 2009), razón por la cual estas propiedades sugieren que LSL<sub>150</sub> actuaría como un potenciador de la solubilidad *in vivo*. En este sentido, recientemente se han descrito en la bibliografía regiones N-terminales de algunas proteínas multidominio que funcionan como verdaderos agentes solubilizantes para sus correspondientes partes C-terminales (Kim *et al.*, 2007).

Teniendo en cuenta su posible función como agente solubilizante natural y que LSL<sub>150</sub> puede ser eficazmente purificada en un único paso cromatográfico de afinidad con matrices de Sepharose® 4B simples (sin modificaciones adicionales), se planteó la hipótesis de trabajo de que LSL<sub>150</sub> podría emplearse como una etiqueta de fusión (etiqueta de afinidad y agente solubilizante) y desarrollar así potenciales aplicaciones biotecnológicas, concretamente en el diseño de protocolos genéricos de expresión de proteínas recombinantes. Asimismo y habiendo comprobado que el módulo lectina N-terminal de la proteína LBL de *L. bicolor* (LBL<sub>152</sub>) se comporta de forma semejante a LSL<sub>150</sub>, la hipótesis planteada se podría extender a la utilización de este módulo como otra potencial etiqueta de fusión.

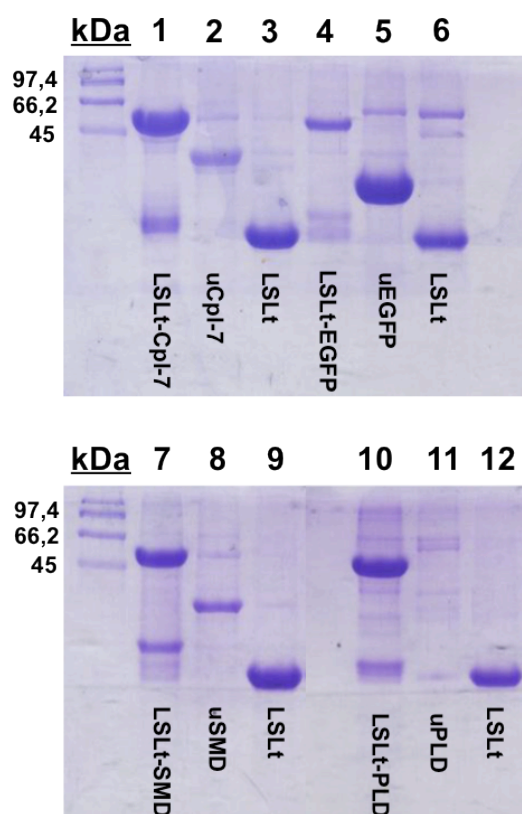
### 9.1. Diseño de un protocolo genérico de expresión-purificación basado en el uso de lectinas

Considerando los dos aspectos citados anteriormente sobre las propiedades de LSL<sub>150</sub> (excelente etiqueta de afinidad y potente agente solubilizante), se llevó a cabo la construcción de un vector de expresión (pKLSLt) diseñado para que contuviera el fragmento génico que codifica al módulo lectina, posicionándolo de tal manera que pudiese fusionarse a proteínas de interés quedando siempre en posición N-terminal. Además, tras la secuencia de LSL<sub>150</sub> se incluyó una región *linker* o conectora (ASSS) y el sitio de reconocimiento y corte (ENLYFQ\*G) de la endoproteasa del virus del grabado del tabaco (TEV), ofreciendo la posibilidad de eliminar la etiqueta de fusión. En paralelo, se siguió la misma estrategia para diseñar dos vectores equivalentes que incluyen al triple mutante LSL<sub>150</sub>- $\alpha$ F (pKELSLt) y a LBL<sub>152</sub> (pKLBLt) en lugar de a LSL<sub>150</sub>. La construcción de los vectores de expresión viene detallada en la sección 7.1 de los *Materiales y Métodos*.

Inicialmente, fueron cuatro las proteínas que se clonaron en el vector pKLSLt para validar a LSL<sub>150</sub> como una posible etiqueta de fusión. Las proteínas elegidas fueron: el módulo catalítico de la endolisina Cpl-7 del virus Cp-7 (pKLSLt-Cpl-7), la proteína verde fluorescente mejorada de *Aequorea victoria* (pKLSLt-EGFP), la esfingomielinasa SMD de *Arcanobacterium haemolyticum* (pKLSLt-SMD) y la fosfolipasa PLD de *Corynebacterium pseudotuberculosis* (pKLSLt-PLD). En todos los casos y de forma resumida, en primer lugar, los extractos celulares clarificados se cargaron en una columna que contenía Sepharose® 4B y que estaba preequilibrada en tampón TSA. Después de un lavado exhaustivo, la fracción retenida se eluyó utilizando tampón TSAL que contiene 0,2 M lactosa. Los resultados obtenidos (**Figura 7.1**) indicaron en todos los casos que las proteínas de fusión de interés se purifican correctamente, aunque siempre aparece una banda correspondiente a la etiqueta de fusión, es decir, LSLt (LSL<sub>150</sub> + *linker* + sitio de reconocimiento TEV), generada de forma espontánea. Después de una digestión en presencia de la endoproteasa TEV y de concentrar la muestra, el digerido se terminó de purificar con una cromatografía de exclusión molecular para retirar la fracción de etiqueta de fusión y eliminar los posibles agregados solubles presentes. Con la excepción de la proteína PLD, que precipitó masivamente después de la digestión con la proteasa, el resto de proteínas se obtuvieron con un grado de pureza muy elevado, tal y como se desprende de los resultados de espectrometría de masas. Definitivamente, la validez de este sistema de purificación se ha demostrado recientemente con las publicaciones de la purificación y cristalización del módulo catalítico de la endolisina



Cpl-7 (Silva-Martín *et al.*, 2010) y también de la 1,3,4,5,6-pentakisfosfato quinasa de *Arabidopsis thaliana*, aunque en este último caso se requirieron más pasos de purificación (Baños-Sanz *et al.*, 2010; González *et al.*, 2010).



**Figura 7.1.** Análisis mediante SDS-PAGE de la producción y purificación recombinante de cuatro proteínas diana fusionadas en el extremo N-terminal a LSL<sub>150</sub>. Todas las proteínas de fusión fueron purificadas en un único paso, empleando columnas de Sepharose® 4B como se describe en *Materiales y Métodos*. Los carriles 1, 4, 7 y 10 se corresponden con las muestras obtenidas de las fracciones purificadas de LSLt-Cpl-7, LSLt-EGFP, LSLt-SMD y LSLt-PLD, respectivamente. Las proteínas de fusión se eluyeron de la Sepharose® 4B lavando con tampón TSAL. Los carriles 2, 5, 8 y 11 se corresponden con las proteínas Cpl-7, EGFP, SMD y PLD, respectivamente, después de un corte en columna con proteasa TEV y un lavado exhaustivo con tampón TSA. Finalmente, los carriles 3, 6, 9 y 12 se corresponden con las muestras eluidas tras la digestión con la proteasa TEV y después de un lavado con tampón TSAL.

Con la intención de minimizar los inconvenientes que aparecían durante la purificación de las anteriores proteínas de fusión, sobre todo en lo referido a la aparición de la etiqueta de fusión por separado y a la limitación de separar proteínas, que después de la digestión tuviesen una masa molecular similar a la de la proteasa o a la de LSLt, se ha diseñado un procedimiento de purificación optimizado de las proteínas de fusión. En primer lugar, la columna de Sepharose® 4B se equilibra en tampón TSA y se carga con el extracto celular clarificado. Después de un lavado exhaustivo con tampón TSA, la columna se equilibra con tampón de digestión (50 mM Tris-HCl pH 8,0, 200 mM NaCl y 0,04% NaN<sub>3</sub>) y posteriormente se añade la proteasa TEV directamente a la columna. La digestión se mantiene a 4 °C con agitación durante toda la noche (~10 h). Una vez finalizada ésta, se lava la columna, que se conecta en serie con otra de afinidad a Ni<sup>2+</sup> (IMAC), en donde la proteasa quedará retenida, pues se utiliza una versión recombinante con una cola de seis histidinas en su extremo N-terminal. Las fracciones no retenidas contienen la proteína de interés pura. Como etapa final, se puede emplear una columna de exclusión molecular para eliminar los posibles

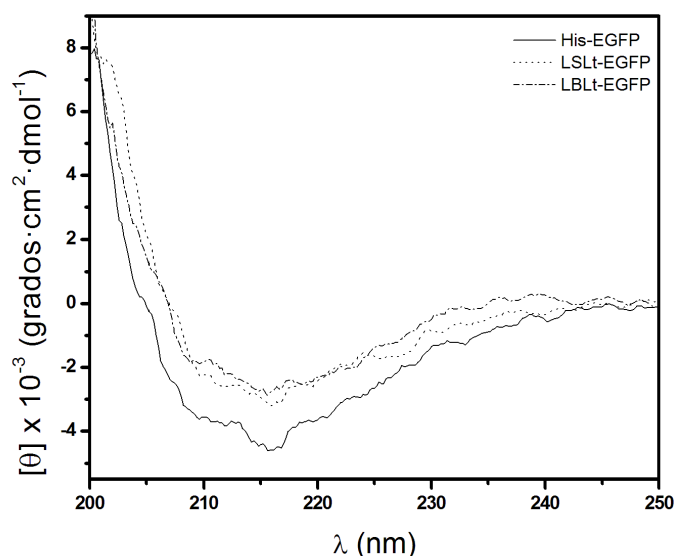
agregados formados. En este procedimiento, además, no se utiliza lactosa, por lo que no es necesario llevar a cabo procesos de diálisis adicionales para eliminar el azúcar.

Los resultados con las cuatro proteínas modelo demostraron que todas se purifican perfectamente con la excepción de la PLD. De hecho, la proteína PLD no llegó a eluir de la columna de afinidad a  $\text{Ni}^{2+}$  (**Figura 7.1**) probablemente debido a la agregación masiva de la proteína diana tras la digestión para quitar LSLt. Éste no es un comportamiento extraño ya que se observa con muchos otros agentes solubilizantes. Por ejemplo, se ha estimado que prácticamente un cuarto de las proteínas expresadas con MBP precipitan cuando se elimina la etiqueta de fusión (Malhotra, 2009). Sin embargo, este hecho no depende de la etiqueta de fusión, puesto que tras su eliminación la solubilidad depende de la propia proteína diana (Nallamsetty y Waugh, 2006). Por este motivo, consideramos en una primera aproximación que el mantenimiento de una proteína soluble cuando ésta está fusionada a la etiqueta de fusión es condición suficiente para clasificar, en este caso, a LSL<sub>150</sub> como un agente solubilizante.

Los otros vectores, pKLBLt y pKELSLt, se comportan de la misma manera que el vector pKLSLt. En ambos casos, se purificó la proteína EGFP con éxito, siguiendo el mismo protocolo de purificación descrito anteriormente. Independientemente del vector utilizado, las muestras finales son siempre fluorescentes, lo que indica que la proteína EGFP está bien plegada, puesto que la fluorescencia depende intrínsecamente del correcto plegamiento de la proteína (Heim *et al.*, 1995). Mediante dicroísmo circular (**Figura 7.2**) se comprobó que el espectro en el UV lejano tanto para la construcción LSLt-EGFP como para LBLt-EGFP es prácticamente el mismo, lo que indica que ambas proteínas de fusión tienen un plegamiento similar. Tanto la proteína EGFP como los módulos lectina tienen un gran porcentaje de estructura tipo  $\beta$  que se refleja en los tres casos con un mínimo cerca de los 217 nm.

Además, se construyó una variante extra del vector de expresión pKLSLt, el vector pKHLSLt, que posee dos etiquetas de fusión en tándem. En este caso, el módulo lectina LSL<sub>150</sub> va acompañado de una cola de seis histidinas en posición N-terminal. La utilización de una segunda etiqueta de fusión suele emplearse para aumentar la pureza con una segunda cromatografía de afinidad (Pryor y Leiting, 1997). Alternativamente, el uso de dos etiquetas de fusión en tándem puede resultar de utilidad cuando una de ellas se emplea en la purificación y la otra se usa como sistema de detección (Lu *et al.*, 1997). Con respecto a cómo poder detectar las proteínas de fusión, cabe destacar que se produjeron anticuerpos policlonales en conejo con actividad anti-LSL<sub>150</sub> y anti-LBL<sub>152</sub>, que permiten llevar a cabo ensayos de inmunodetección de las proteínas de fusión que incorporan estas etiquetas. En

el caso específico de las proteínas producidas en el vector pKHLSLt también se podrían utilizar anticuerpos monoclonales anti-*His tag*, disponibles comercialmente.



**Figura 7.2.** Espectros de dicroísmo circular en la región del UV lejano de His-EGFP (línea continua) y las construcciones LSLt-EGFP (línea de puntos) y LBLt-EGFP (línea de rayas y puntos) en tampón TSA. Los espectros fueron medidos a 25 °C y con una cubeta de 0,1 cm de paso óptico.

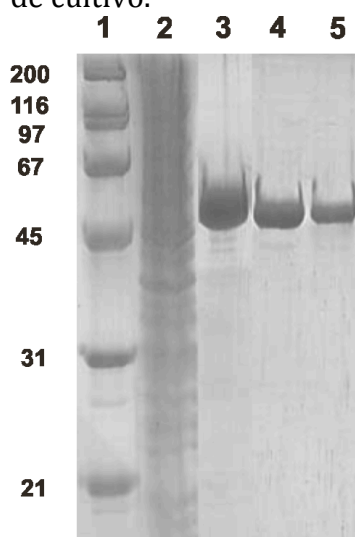
## 9.2. LSLt-Bgl: descripción detallada de un ejemplo práctico

### 9.2.1. Antecedentes: la $\beta$ -galactosidasa de *Lactobacillus plantarum*, una enzima difícil de purificar

Una de las líneas de investigación activas del Grupo de Cristalografía y Biología Estructural (CBE), en colaboración con el Grupo de Biotecnología Bacteriana del Instituto de Ciencia y Tecnología de Alimentos y Nutrición (ICTAN, CSIC), está enfocada en el estudio de enzimas de interés industrial procedentes de la bacteria láctica *Lactobacillus plantarum*. Hoy en día *L. plantarum* resulta un microorganismo de interés biotecnológico ya que se utiliza con relativa frecuencia en procesos fermentativos dentro de la industria alimentaria. Por ejemplo, es importante su contribución en el proceso de vinificación debido a que lleva a cabo la fermentación maloláctica (Olsen *et al.*, 1991). Gracias a la secuenciación del genoma de este microorganismo, se ha facilitado enormemente la identificación a nivel génico de genes potenciales que expresarían proteínas de interés en Biotecnología. En particular, uno de estos genes, el *lp\_3629*, codificaría una potencial glucosidasa debido al alto porcentaje de similitud que presenta con otras glucosidasas de bacterias lácticas (Spano *et al.*, 2005). Las glucosidasas, de hecho, son enzimas muy interesantes tanto en procesos de síntesis de carbohidratos, puesto que permiten obtener determinados compuestos de un modo regio y estereoespecífico (Bojarova y Kren, 2009), como en numerosas aplicaciones biotecnológicas (Bhatia *et al.*, 2002), tales como la detoxificación de alimentos, la conversión

de biomasa y la potenciación de los aromas en bebidas. En este contexto, se decidió llevar a cabo un estudio estructural profundo de esta enzima, denominada Bgl.

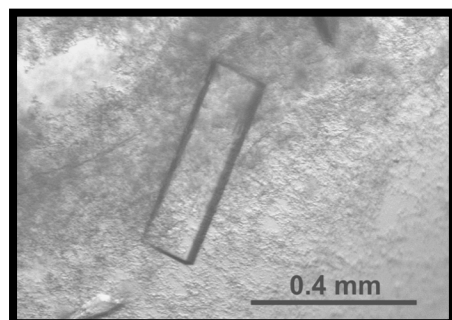
El gen *lp\_3629* fue clonado a partir del ADN genómico de *L. plantarum* CECT 748<sup>T</sup> en el vector pURI3 (de las Rivas *et al.*, 2007), incluyendo en el extremo N-terminal una etiqueta de afinidad y un sitio de corte por enteroquinasa (MGGSHHHHHHGDDDDK<sup>\*</sup>M). La expresión de la proteína Bgl se llevó a cabo en células de *E. coli* BL21(DE3), induciendo con IPTG (0,4 mM) y dejando los cultivos crecer durante ~20 horas a 22 °C. Después de la sobreexpresión, las células fueron recogidas y los precipitados se resuspendieron en tampón TSA en presencia de inhibidores de proteasas. Tras la lisis de las células, el extracto celular clarificado se purificó en un primer paso mediante una cromatografía de afinidad a Ni<sup>2+</sup>, utilizando un gradiente lineal de imidazol con las soluciones TSIA-A y B. El siguiente paso de purificación consistió en una cromatografía de intercambio iónico, utilizando una columna HiTrap Q HP (GE Healthcare Life Sciences) y un gradiente lineal de NaCl (10-500 mM) en tampón 20 mM Tris-HCl pH 8,0. Las fracciones que contenían la proteína Bgl se dializaron frente a tampón TSA con 2 mM DTT (TSA-reducido) y, posteriormente, después de concentrarse, la proteína se cargó en una columna de exclusión molecular, donde se llevó a cabo el último paso de purificación. El proceso completo de purificación de Bgl se siguió electroforéticamente (**Figura 7.3**) y su producción final se estimó en 15 mg/L de cultivo.



**Figura 7.3.** SDS-PAGE (10% de poliacrilamida) de la purificación de la enzima Bgl. En el carril 1 se muestran los patrones de masa molecular Broad Range (Bio-Rad); en el carril 2, la fracción soluble del extracto celular; en el carril 3, la fracción eluída de la cromatografía de afinidad a Ni<sup>2+</sup>; en el carril 4, la fracción eluída de la cromatografía de intercambio iónico; y en el carril 5, la fracción final eluída de la cromatografía de exclusión molecular.

Una vez se hubo purificado la enzima, se llevaron a cabo los ensayos de cristalización. En un primer lugar, aparecieron cristales en muchas condiciones que tenían en común el PEG 3350 como agente precipitante. Después de un proceso lento de optimización, los mejores cristales se obtuvieron a una concentración de enzima de 9 mg/ml en 15% (*p/v*) PEG 3350, 0,2 M fosfato dibásico de amonio, 0,1 M cacodilato de sodio pH 6,4, 2 mM DTT y

detergente CTAB como aditivo (**Figura 7.4**). Estos cristales fueron medidos en la línea BM16 del ESRF (Grenoble, Francia) a 100 K y utilizando una longitud de onda de 0,979 Å. Previamente, los cristales fueron crioprotegidos en una solución con un 25% (v/v) de 2-metil-2,4-pentanodiol (MPD). Con el mejor set de datos se alcanzó un 94% de completitud y una resolución de hasta 2,48 Å. El indexado y posterior escalado de los datos reveló que los cristales poseen un grupo espacial *C2* con una celda unidad de  $a = 196,7$  Å,  $b = 191,7$  Å,  $c = 105,9$  Å y  $\alpha = \gamma = 90^\circ$  y  $\beta = 102,74^\circ$  (Acebrón *et al.*, 2009). El resto de datos del procesamiento quedan recogidos en la **Tabla 7.1**.



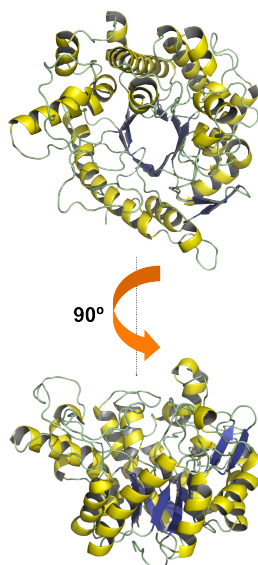
**Figura 7.4.** Fotografía de un cristal de la proteína Bgl recombinante de *L. plantarum* (9 mg/mL) crecido a 18 °C en 15% (p/v) PEG 3350, 0,2 M fosfato dibásico de amonio, 0,1 M cacodilato de sodio pH 6,4, 2 mM DTT y detergente CTAB como aditivo mediante el método de difusión de vapor en gota colgante.

Parámetros	Bgl	Bgl2
Recogida de datos y procesamiento		
Grupo espacial	<i>C2</i>	<i>C2</i>
Moléculas en la UA	6	6
Celda unidad (Å)	196,7, 191,7, 105,9 ( $\beta = 102,7^\circ$ )	198,7, 192,1, 105,8 ( $\beta = 102,5^\circ$ )
Longitud de onda (Å)	0,979	1,067
Resolución (Å)	67,7–2,48	53,4–2,60
Reflexiones totales	375.701	331.239
Reflexiones únicas	126.809	112.830
Compleitud	94,0 (94,0)	95,0 (89,9)
Multiplicidad	3,0 (2,9)	2,9 (2,7)
$R_{merge}$ (%)	11,5 (43,2)	9,7 (44,6)
$I/\sigma(I)$	9,7 (2,2)	8,1 (1,7)
Refinamiento		
$R_{work}/R_{free}$	0,183/0,239	0,179/0,220
RMSD de los enlaces (Å)	0,020	0,008
RMSD de los ángulos (°)	1,840	1,080
Factor <i>B</i> medio (Å <sup>2</sup> )	9,47	47,90
Número de átomos		
Proteína	22.377	22.374
Fosfato	60	---
Agua	849	385
Mapa de Ramachandran		
Favorecidos (%)	96,2	96,0
Permitidos (%)	3,6	4,0
Prohibidos (%)	0,2	0,0

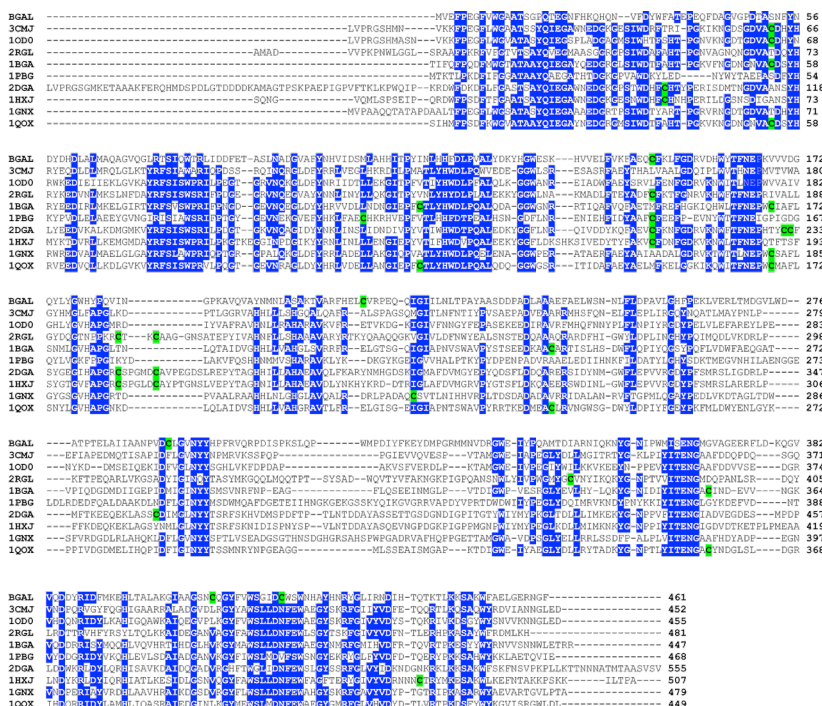
\*Los valores de la concha de mayor resolución se muestran en paréntesis

**Tabla 7.1.** Datos de difracción y estadísticas del refinamiento de la estructuras de la proteína Bgl recombinante y su respectivo mutante Bgl2 (Cys211Ser, Cys292Ser).

La estructura de Bgl se resolvió mediante reemplazo molecular con PHASER (McCoy *et al.*, 2007), utilizando como modelo de reemplazo la estructura de la  $\beta$ -glucosidasa mejorada de un metagenoma del suelo (PDB 3CMJ) (Nam *et al.*, 2008). El plegamiento de Bgl (**Figura 7.5**) se caracteriza por la presencia de un barril  $(\beta\alpha)_8$  o barril TIM, que contiene ocho hebras  $\beta$  paralelas entre sí que forman un pequeño cilindro. Los residuos de las conexiones entre las hebras  $\beta$  forman un total de ocho hélices  $\alpha$  intercaladas entre las hebras que se colocan externamente a ellas, formando una envoltura. Este plegamiento es típico de la familia 1 de las glicosil hidrolasas (GH1), grupo en el que algunas de las proteínas comparten hasta un 30% de identidad de secuencia con Bgl (**Figura 7.6**).



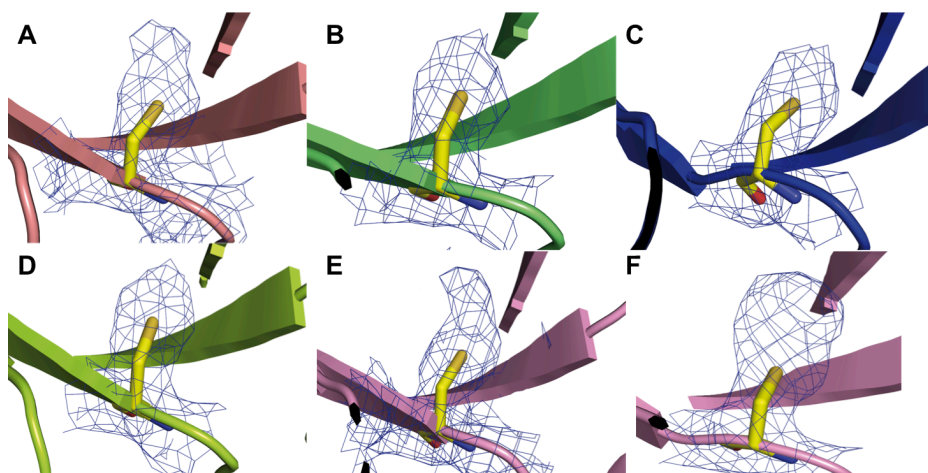
**Figura 7.5.** Visión cenital (arriba) y frontal (abajo) de la estructura de la Bgl recombinante de *L. plantarum*. En el modelo de *bucles y cintas* se muestra el barril ( $\beta\alpha$ )<sub>8</sub>, típico de la familia 1 de las glicosil hidrolasas (GH-1), representando las hélices  $\alpha$  en *amarillo* y las hebras  $\beta$  en *azul*.



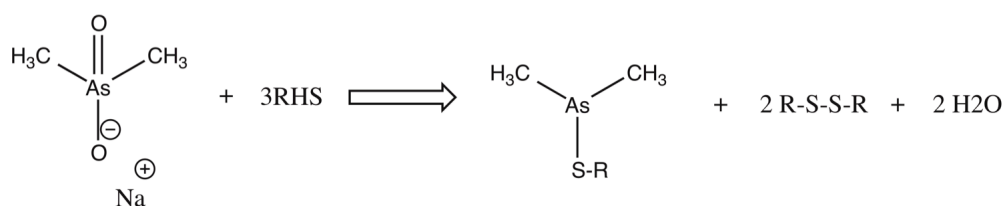
**Figura 7.6.** Alineamiento de secuencia múltiple de Bgl con algunas glicosidasas con una estructura tridimensional conocida que muestran similitud de secuencia con Bgl según BLAST. Las posiciones que están altamente conservadas se representan en azul y los residuos de Cys en verde. BGAL, *L. plantarum* Bgl; 3CMJ,  $\beta$ -glucosidasa de un metagenoma del suelo; 10D0,  $\beta$ -glucosidasa de *Thermogota marítima*; 2RGL,  $\beta$ -glucosidasa del arroz; 1BGA,  $\beta$ -glucosidasa de *Bacillus plumyxa*; 1PBG, 6-fosfo- $\beta$ -galactosidasa de *Lactococcus lactis*; 2DGA,  $\beta$ -glucosidasa del trigo; 1HXJ,  $\beta$ -glucosidasa del maíz; 1GNX,  $\beta$ -glucosidasa de *Streptomyces sp*; 1QOX,  $\beta$ -glucosidasa de *Bacillus circulans sp. alkalophilus*.



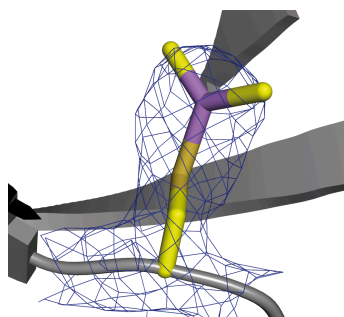
El análisis en detalle de la estructura de Bgl reveló la presencia de una densidad electrónica extra alrededor de la Cys-292 (expuesta en la superficie de la proteína) en tres de las seis moléculas independientes de la unidad asimétrica (**Figura 7.7**). En este sentido y atendiendo a la condición de cristalización, se observó que en otras proteínas resueltas anteriormente había sido descrita una reacción química en la que participan como reactivos el cacodilato de sodio, el DTT y grupos sulfhidrilo de las proteínas (R-SH) (Scott *et al.*, 1993) (**Figura 7.8**). Teniendo en cuenta que el cacodilato y el DTT están en la condición de cristalización, ciertos residuos de cisteína localizados en la superficie podrían ser modificados covalentemente. De hecho, la incorporación de la modificación química al modelo estructural de Bgl se ajustó perfectamente a la densidad electrónica observada (**Figura 7.9**). Por lo tanto, quedaba demostrado que esta modificación es consistente con las condiciones de cristalización y justificaría perfectamente la densidad electrónica observada alrededor de algunos residuos de Cys-292. La susceptibilidad a que los residuos de Cys-292 sean modificados químicamente podría explicar los problemas de solubilidad que tuvieron lugar durante la purificación de la proteína, que obligaron a incorporar 2 mM DTT al tampón final. En ocasiones, estos problemas se acrecentaron de tal manera que resultó imposible resolubilizar la proteína y proseguir con la purificación.



**Figura 7.7.** Vista detallada de las cadenas laterales de Cys-292 en cada una de las moléculas (A-F) de la unidad asimétrica del cristal de Bgl. El mapa de densidad electrónica  $2F_o - F_c$  (azul) a  $1\sigma$  revela una densidad extra alrededor del residuo de Cys-292 en las moléculas A, E y F.



**Figura 7.8.** Reacción química en la que se muestra cómo una molécula de cacodilato de sodio participa junto con tres grupos sulfhidrilo en la producción de un derivado con un enlace covalente As-S entre el cacodilato de sodio y uno de los grupos sulfhidrilo disponibles.



**Figura 7.9.** Residuo de Cys-292 de la proteína Bgl modificado químicamente por acción del ion cacodilato. La densidad electrónica del mapa  $2F_o - F_c$  (azul) a  $1\sigma$  es consistente con la presencia de un átomo de arsénico (As) junto con los dos grupos metilo.

### 9.2.2. Resolviendo los problemas de solubilidad de Bgl: ingeniería de proteínas vs proteínas de fusión

Con la intención de proseguir con el estudio de la enzima y obtener complejos para entender su mecanismo de acción, se plantearon dos estrategias distintas para optimizar la purificación y evitar los problemas de solubilidad que presentaba la versión anterior: (1) ingeniería de proteínas y (2) expresión-purificación con el sistema pKLSLt.

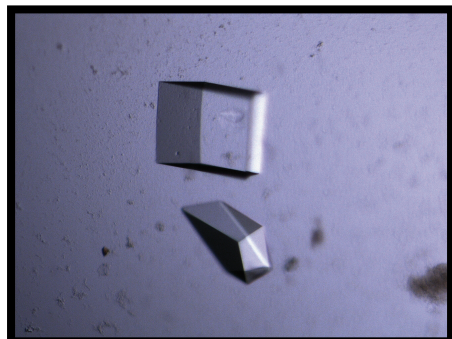
#### **Alternativa 1:** Ingeniería de proteínas

Teniendo en cuenta la primera estrategia, se diseñó un doble mutante de la enzima Bgl en el que los aminoácidos Cys-211 y Cys-292 fueron sustituidos por residuos de Ser (Bgl2). Aunque la Cys-292 es la única que presentaba modificaciones covalentes en la estructura cristalina de la Bgl, ambos residuos de cisteína se encuentran expuestos en la superficie de la proteína, por lo que podrían estar implicados igualmente en el fenómeno de agregación. El doble mutante se purificó siguiendo el mismo protocolo de la proteína silvestre, no observándose ningún problema de solubilidad, razón por la cual no fue necesario añadir DTT en todo el proceso. Este comportamiento explica, por tanto, que posiblemente uno de los dos o los dos residuos de Cys eran los responsables de su susceptibilidad a agregar durante la purificación de la proteína silvestre. A pesar de no tener problemas con la solubilidad de la proteína, el rendimiento disminuyó drásticamente, obteniéndose una producción de tan solo ~2 mg/L de cultivo.

El mutante Bgl2 cristalizó fácilmente, obteniéndose los mejores cristales a una concentración de 12,5 mg/mL en la condición de cristalización 20% (*p/v*) PEG 3350, 0,2 M citrato de sodio y 50 mM espermidina (**Figura 7.10**). Estos cristales, crioprotegidos en un 25% (*v/v*) de glicerol, fueron medidos en la línea ID23-1 del ESRF (Grenoble, Francia). El procesamiento de los datos de difracción demostró que los cristales de Bgl2 eran esencial-

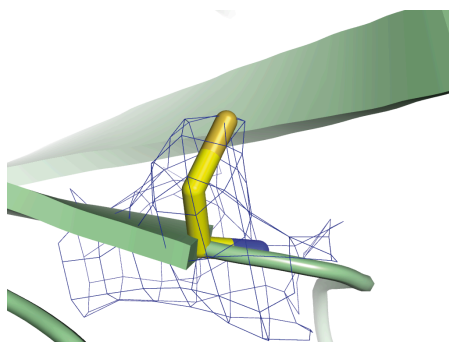


mente los mismos que los de la proteína silvestre, dado que ni el grupo espacial ni los parámetros de la celda unidad apenas cambiaron. El resto de estadísticas del procesamiento está resumido en la **Tabla 7.1**.



**Figura 7.10.** Cristales de Bgl2 (Cys211Ser, Cys292Ser) crecidos a 18 °C en 20% (p/v) PEG 3350, 0,2 M citrato de sodio y 50 mM espermidina.

La resolución estructural del mutante Bgl2 se llevó a cabo mediante reemplazo molecular, usando como modelo de reemplazo la estructura de la proteína silvestre. A la vista de la estructura, el plegamiento global de la proteína se mantiene intacto con la diferencia de que no aparecen las densidades electrónicas que se apreciaban anteriormente en torno a las cadenas laterales de los residuos de Cys-292 en Bgl (**Figura 7.11**).



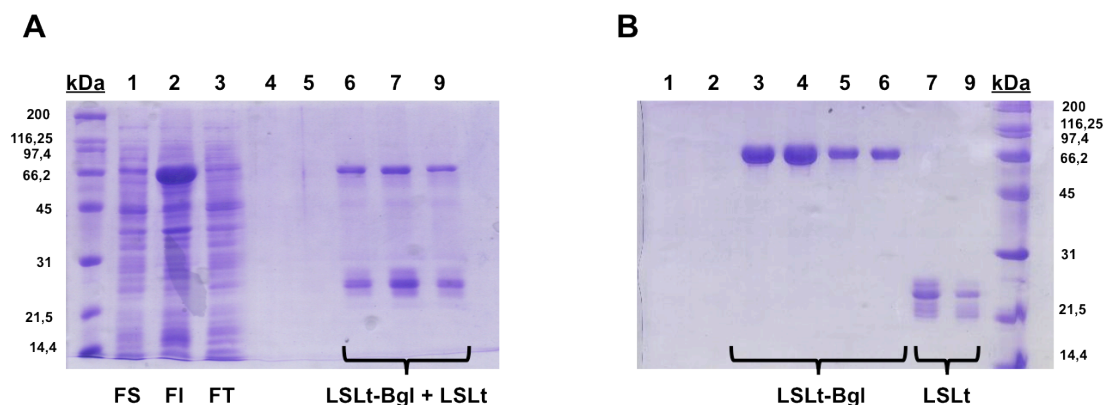
**Figura 7.11.** Densidad electrónica del mapa  $2F_o - F_c$  (azul) a  $1\sigma$  alrededor de la cadena lateral de la Ser-292 de Bgl2. A diferencia de lo que se veía en la estructura de Bgl, no se observa ninguna densidad electrónica extra alrededor de los residuos de Ser-292 en ninguna de las moléculas de la unidad asimétrica.

### **Alternativa 2:** Expresión-purificación con el sistema pKLSLt

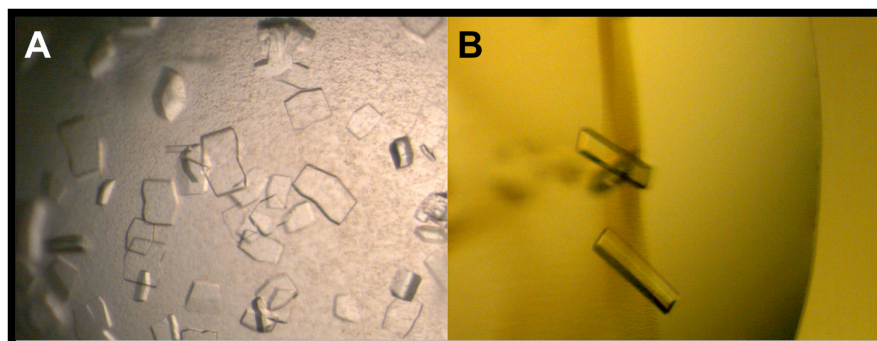
Aunque la producción del mutante solventó los problemas de solubilidad durante la purificación, el rendimiento de producción disminuyó drásticamente. Por este motivo, también se llevó a cabo la otra alternativa que se planteó para purificar Bgl, consistente en utilizar el vector pKLSLt para expresar y aumentar así la solubilidad de la enzima. De esta manera, además, la enzima se obtendría sin ningún cambio a nivel de la secuencia de aminoácidos a diferencia de lo que ocurre con Bgl2.

La expresión y la purificación de la proteína LSLt-Bgl se llevaron a cabo siguiendo el mismo protocolo aplicado en el resto de proteínas de fusión. La purificación de la proteína

de fusión simplifica sobremanera el proceso seguido hasta el momento porque simplemente consta de dos sencillos pasos, la cromatografía de afinidad a Sepharose® 4B y la cromatografía de exclusión por tamaño. En la **Figura 7.12** se muestran los geles del seguimiento de la purificación, donde se aprecia el nivel de pureza obtenido tras la última etapa. El rendimiento de producción determinado (~7 mg/L de cultivo) fue superior al obtenido para Bgl2, pero continúa siendo inferior al de Bgl, lo que indicaría que sobreexpresar la proteína de fusión es más costoso para la célula. Finalmente, la etiqueta de fusión LSLt fue eliminada con éxito mediante una digestión con la proteasa TEV, obteniéndose la proteína uBgl (versión sin la etiqueta de fusión que posee el residuo de Gly en posición N-terminal que permanece tras la digestión). Sorprendentemente, uBgl se mantuvo soluble en el tampón TSA, sin mostrar signos de precipitación, por lo que tampoco fue necesario añadir ningún agente reductor.

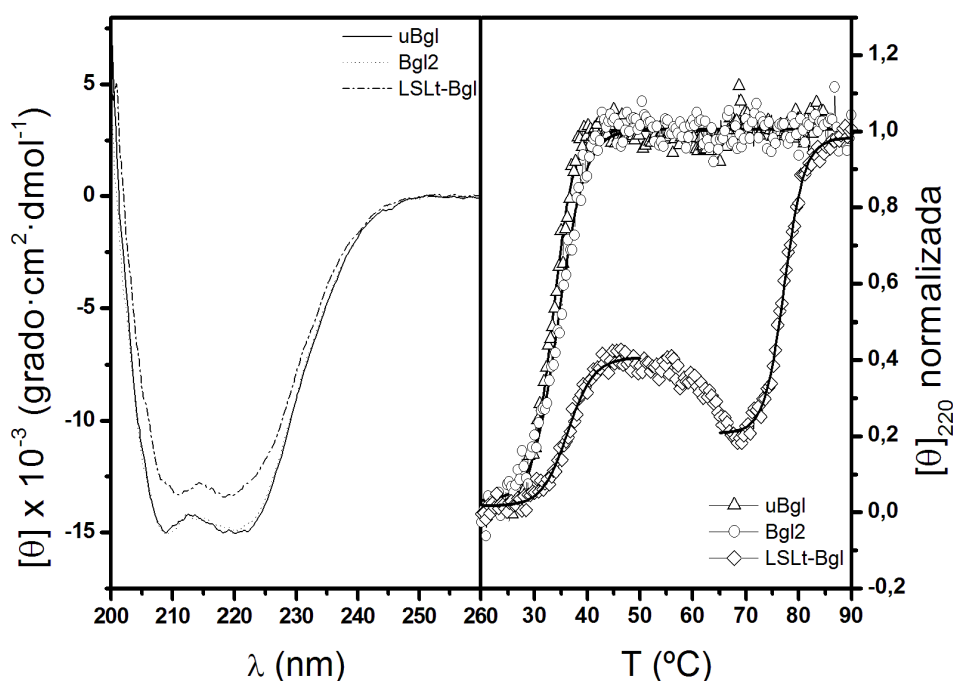


**Figura 7.12.** SDS-PAGE (12% de poliacrilamida) de la purificación de LSLt-Bgl. En **A** se muestra el gel del primer paso de purificación mediante la cromatografía de afinidad a Sepharose® 4B, donde FS es la fracción soluble del extracto celular, FI es la fracción insoluble y FT es la fracción no retenida. En **B** se muestra el último paso de purificación, la cromatografía de exclusión molecular, que permite separar la proteína de fusión LSLt-Bgl de la etiqueta de afinidad generada espontáneamente (LSLt).



**Figura 7.13.** Cristales de uBgl obtenidos en diferentes condiciones de cristalización a 18 °C mediante difusión de vapor. Los cristales de **A** crecieron en 20% (p/v) PEG 3350 y 0,2 M formiato de sodio. Los cristales de **B**, en cambio, aparecieron en 25% (p/v) PEG 3350, 0,1 M HEPES pH 7,5 y 0,2 M acetato de amonio.

La resolución estructural de uBgl está aún en curso porque, a pesar de haber cristalizado en otras condiciones distintas a las iniciales (**Figura 7.13**), los patrones de difracción obtenidos no son de suficiente calidad para llevar a cabo el análisis cristalográfico. Por esta razón, se realizaron estudios estructurales mediante dicroísmo circular con el fin de comprobar si, al menos, la estructura secundaria se sigue manteniendo. A la vista de los espectros (**Figura 7.14**), se desprende que uBgl posee la misma estructura, rica en hebra  $\beta$ , que presenta el mutante Bgl2, que, en definitiva, es la misma que poseía la primera versión de la proteína Bgl recombinante, tal y como se desprende de lo visto en las estructuras cristalográficas. En el caso de LSLt-Bgl, el espectro presenta en general la misma forma que uBgl y Bgl2, pero debido a la contribución del módulo lectina, la señal de dicroísmo es más débil.



**Figura 7.14.** Dicroísmo circular en el UV lejano de Bgl2, LSLt-Bgl y uBgl. A la izquierda se muestran los espectros de DC de uBgl (línea continua), Bgl2 (línea punteada) y LSLt-Bgl (línea de puntos y rayas). Los espectros fueron medidos a 25  $^{\circ}\text{C}$  y con una cubeta de 0,1 cm de paso óptico. A la derecha se representan los perfiles de desnaturalización térmica de uBgl (triángulos), Bgl2 (círculos) y LSLt-Bgl (rombos). La línea continua se corresponde al ajuste a una curva sigmoidea de los datos experimentales para la determinación de la temperatura media de transición ( $T_m$ ).

Por otro lado, se estudió la estabilidad térmica de las tres proteínas: Bgl2, uBgl y LSLt-Bgl (**Figura 7.14**). Como era de esperar, debido a que la estructura global de la proteína no cambia, la temperatura de transición media de Bgl2 y uBgl es prácticamente la misma,  $33,7 \pm 0,1$  y  $34,6 \pm 0,1$   $^{\circ}\text{C}$ , respectivamente. Por otro lado, el seguimiento de la desnaturalización térmica de LSLt-Bgl mediante DC revela su estructura modular, observándo-

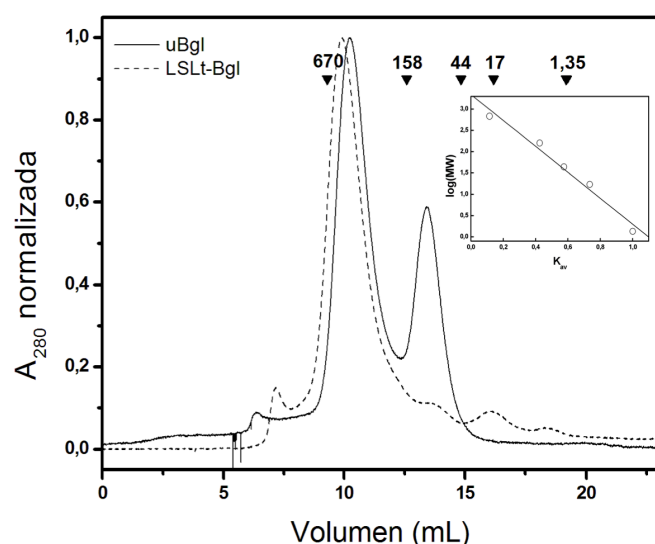
se una primera transición a los  $36,6 \pm 0,2$  °C, correspondiente a la desnaturalización de Bgl, seguida de otra, a los  $77,3 \pm 0,1$  °C, que es similar a la identificada previamente para el módulo lectina. Este resultado indica que ambas proteínas se comportan como módulos independientes. Por último, en los tres resultados la temperatura de transición asociada a Bgl es muy baja, lo cual podría influir también en los problemas de agregación y precipitación que presenta. En este sentido, se ha visto que los cambios de pH influyen considerablemente en la estabilidad térmica, ya que si las medidas se hacen a pH 7,0 (una unidad de pH por debajo de la del tampón TSA), la estabilidad aumenta alrededor de 5–6 °C. Este resultado indica, por tanto, la importancia de la elección del tampón que se va a ser utilizado durante el proceso de purificación y los posteriores ensayos de caracterización de la proteína (Structural Genomics *et al.*, 2008).

### 9.2.3. La formación del oligómero de Bgl no se ve afectada por la etiqueta de fusión

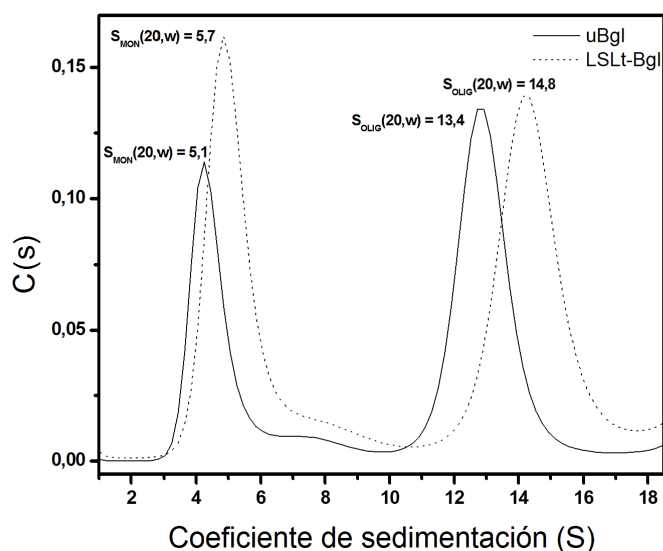
El estado oligomérico de la proteína Bgl había sido estudiado con anterioridad mediante cromatografía de exclusión molecular, observándose la coexistencia de dos especies consistentes con las masas moleculares de un monómero ( $70,0 \pm 3,4$  kDa) y de un pentámero/hexámero ( $390,6 \pm 4,3$  kDa) (Acebrón *et al.*, 2009). Este mismo experimento se llevó a cabo con uBgl para comprobar si se seguía observando la formación de una especie oligomérica. Los resultados (**Figura 7.15**) indican que uBgl existe en solución como una mezcla de dos especies de  $41,6 \pm 1,2$  kDa y  $301 \pm 23$  kDa, correspondientes al monómero (53 kDa) y a un posible hexámero, respectivamente. Este comportamiento coincide, por tanto, con el descrito previamente para la proteína Bgl, lo que indicaría que la enzima se presenta siempre en solución como un sistema asociativo entre una especie monomérica y un ensamblado molecular que probablemente contenga seis unidades. El perfil de elución también demuestra que la capacidad de oligomerizar no se pierde cuando la proteína permanece unida a la lectina de la etiqueta de fusión. De hecho, a la vista de la **Figura 7.15**, parece que el equilibrio de asociación estuviese completamente desplazado a la formación del oligómero.

A pesar de que los resultados de la cromatografía analítica indican que la enzima posee estructura cuaternaria, el estudio del empaquetamiento cristalino de los cristales de Bgl con el servidor PISA (Krissinel y Henrick, 2007) indica la presencia de monómeros. Las condiciones de cristalización podrían modificar el equilibrio asociativo y desplazarlo hacia

la forma monomérica, más susceptible de cristalizar. Para comprender mejor el estado oligomérico, se llevaron a cabo experimentos de ultracentrifugación analítica. Los perfiles de distribución de los coeficientes de sedimentación [ $s(20,w)$ ] (**Figura 7.16**) de las proteínas uBgl y LSLt-Bgl corroboraron la existencia de las dos especies en solución. En ambos casos, las dos proteínas presentan coeficientes de sedimentación parecidos, siendo ligeramente mayores en el caso de LSLt-Bgl debido a su mayor masa molecular (71,5 kDa). Los resultados de equilibrio de sedimentación también confirmaron la formación de un equilibrio monómero-hexámero, que tiene lugar en las dos proteínas analizadas. Todos estos resultados indican, por lo tanto, que el módulo lectina no interfiere en los mecanismos de auto-asociación. Este hecho es relevante en sí mismo, ya que si la actividad biológica es dependiente del equilibrio asociativo, la versión LSLt-Bgl podría emplearse para estudios acerca de la función de la enzima.



**Figura 7.15.** Cromatografía analítica de exclusión molecular de uBgl y LSLt-Bgl. Los dos perfiles de elución se muestran conjuntamente con las posiciones de los picos de elución de los patrones de masa molecular conocida en kDa (*triángulos invertidos*). En el interior del gráfico se muestra la recta de calibrado ( $\log(MW)$  vs  $K_{av}$ ) determinada para calcular la masa molecular de las diferentes especies.



**Figura 7.16.** Perfiles de distribución de los coeficientes de sedimentación [ $s(20,w)$ ] de LSLt-Bgl (*línea discontinua*) y uBgl (*línea continua*). Las medidas se realizaron a 20 °C y 45.000 rpm en tampón TSA.

#### 9.2.4. Estudio cualitativo de la actividad catalítica de la enzima

A pesar de la extensa descripción que hay acerca de las actividades glicosidasas en bacterias lácticas, las correspondientes enzimas han sido escasamente caracterizadas. Debido a los análisis de homología de secuencia, la enzima Bgl, está clasificada como una potencial  $\beta$ -glucosidasa perteneciente a la familia 1 de las glicosil hidrolasas (GH1). Esta familia engloba un gran número de especificidades, incluyendo  $\beta$ -glucosidasas (EC 3.2.1.21),  $\beta$ -galactosidasas (EC 3.2.1.23) y 6-fosfo- $\beta$ -galactosidasas (EC 3.2.1.85), entre otras. Con la intención de demostrar bioquímicamente que Bgl es realmente una  $\beta$ -glucosidasa, se llevaron a cabo ensayos de actividad enzimática frente al sustrato *p*-nitrofenil- $\beta$ -D-glucósido (pNPG). Sorprendentemente, este compuesto no fue hidrolizado por ninguna de las muestras de proteína estudiadas (Bgl, Bgl2, uBgl y LSLt-Bgl) (**Tabla 7.2**). Teniendo en cuenta que la familia GH1 contiene otras actividades enzimáticas, se llevaron a cabo ensayos de actividad para determinar si Bgl tiene actividad  $\beta$ -galactosidasa y/o 6-fosfo- $\beta$ -galactosidasa. Las cuatro versiones de la enzima Bgl catalizaron la liberación de  $\beta$ -galactosa y *p*-nitrofenol con la consecuente aparición de una coloración amarilla. El sustrato *o*-nitrofenil- $\beta$ -D-galactósido (oNPG) mostró una coloración más intensa que el *p*-nitrofenil- $\beta$ -D-galactósido (pNPGal), lo que indica que la posición del sustituyente en el anillo es importante. Por otro lado, el sustrato *o*-nitrofenil- $\beta$ -D-galactósido-6-fosfato (oNPG6P) no se hidrolizó, por lo que en principio Bgl sólo es específica de  $\beta$ -D-galactósidos.

Cabe destacar que la actividad de la enzima Bgl recombinante es mayor en presencia de DTT que sin el agente reductor. Probablemente, este resultado esté relacionado con la tendencia a agregar de la enzima cuando no hay DTT en el medio. Es curioso cómo, sin embargo, el mutante Bgl2 presenta muy poca actividad. El patrón de distribución de los residuos de Cys en las proteínas homólogas a Bgl no está conservado, por lo que cabría esperar que las Cys no estuvieran implicadas en la catálisis, y de hecho, no se encuentran cerca del centro activo. En cambio, la pérdida de actividad es evidente tanto en presencia como en ausencia de DTT. Por otro lado, tanto uBgl como LSLt-Bgl mostraron la mayor actividad (mayor coloración) incluso en ausencia de DTT. En ambos casos, parece ser que la incorporación de la etiqueta de fusión, al eliminar los problemas de agregación por completo, favorecen a la actividad enzimática.

La identificación de Bgl como una nueva  $\beta$ -galactosidasa, y no como una  $\beta$ -glucosidasa, como se había predicho, indica que la adquisición de nuevas especificidades

por los miembros de la familia GH1 es un mecanismo evolutivo de adaptación de estas enzimas (Henrissat, 1991). También hay que indicar que este resultado demuestra que las asignaciones de función basadas en datos genómicos deben verificarse siempre en último término experimentalmente con la enzima aislada (White, 2006).

	Bgl		Bgl2		LSLt-Bgl		uBgl	
	- 1mM DTT	+ 1mM DTT	- 1mM DTT	+ 1mM DTT	- 1mM DTT	+ 1mM DTT	- 1mM DTT	+ 1mM DTT
pNPG	–	–	–	–	–	–	–	–
oNPG6P	–	–	–	–	–	–	–	–
pNPGal	++	++	+	+	++	++	++	++
oNPG	++	+++	+	+	+++	+++	+++	+++

**Tabla 7.2.** Estudio de la actividad enzimática de la enzima Bgl de *L. plantarum* frente a diferentes sustratos. La actividad enzimática se siguió colorimétricamente. El color gris indica la ausencia de actividad, mientras que las diferentes tonalidades de amarillo indican la presencia de actividad. La intensidad (+) del color está relacionada con la capacidad hidrolítica de la enzima. El estudio se llevó a 37 °C y en un intervalo de tiempo de 10 minutos.

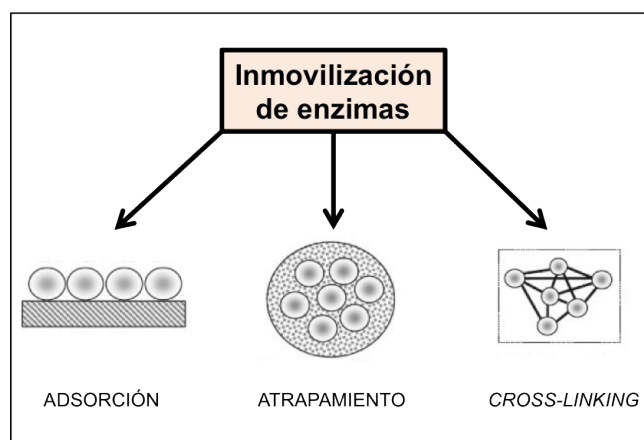
## 10. Inmovilización de biomoléculas en matrices de agarosa

El desarrollo de los procesos de inmovilización de enzimas ha permitido llevar a cabo gran cantidad de biotransformaciones de interés a nivel industrial. Las principales razones por las que se inmovilizan enzimas surgen de la necesidad de mejorar la estabilidad de éstas y de su reutilización, puesto que aumenta en comparación con el uso de enzimas libres. El uso de enzimas en la síntesis de productos químicos solventa algunos de los problemas de la síntesis química tradicional, como la protección y/o activación de grupos funcionales determinados, haciendo la producción más económica, generando menos residuos y siendo más sostenible desde un punto de vista energético (Sheldon y van Pelt, 2013). Además, la inmovilización reduce drásticamente la presencia de contaminantes proteicos en el producto final, evitando posibles problemas de alergia. Aun con todo, el uso de biocatalizadores acarrea nuevos problemas como el mantenimiento de la estabilidad de la enzima durante tiempos largos o la producción de éstas a gran escala.

En la actualidad existen básicamente tres métodos para llevar a cabo la inmovilización de enzimas: (i) unión a soportes, (ii) encapsulación o atrapamiento y (iii) entrecruzamiento o *cross-linking* (**Figura 7.17**). Aunque todas estas metodologías se usan en la industria, la inmovilización de proteínas específica a soportes que llevan una etiqueta de fusión en su extremo N o C-terminal ha empezado a usarse en muchas aplicaciones biotecnológicas (Saleemuddin, 1999; Terpe, 2003). Los problemas asociados a la inmovilización a soportes están relacionados principalmente con el desprendimiento de la enzima del soporte (*leaching*) bajo determinadas condiciones de reacción o debido a las altas concentraciones



de producto. Para solventar este problema, se han desarrollado métodos de fijación covalente de la enzima, que muestran algunas desventajas como la pérdida aparente de actividad (ocultamiento del sitio activo) o el elevado coste (Sheldon y van Pelt, 2013).



**Figura 7.17.** Diferentes métodos de inmovilización de enzimas

Las propiedades de las enzimas inmovilizadas a soportes no solo dependen de la enzima en cuestión, sino del material usado como soporte. Éstos pueden ser de diferente naturaleza: (a) polímeros orgánicos sintéticos, (b) biopolímeros y (c) polímeros inorgánicos. A pesar de que los polímeros orgánicos e inorgánicos son los más utilizados debido a que se pueden derivatizar para llevar a cabo el anclaje covalente de las enzimas, los biopolímeros constituyen una alternativa más económica.

La incorporación de etiquetas de fusión en el complejo mundo de las inmovilizaciones enzimáticas no supone una novedad. Sin embargo, el uso de lectinas con tal fin, no es una estrategia frecuente. Los dominios lectina se han usado ampliamente como biosensores y en las modernas tecnologías de los *microarray* (Hu y Wong, 2009; Tetala y van Beek, 2010). Por ejemplo, la ConA se ha usado para la inmovilización de glicoproteínas (Tang *et al.*, 2010). No obstante, tales dominios lectina se usan anclados a una matriz sólida y no como una etiqueta de fusión para la producción recombinante de proteínas.

La inmovilización de enzimas continúa siendo un campo de interés, tanto para la industria como para el mundo académico. La tendencia hoy en día es la búsqueda de formas relativamente sencillas de inmovilizar los biocatalizadores, no requiriendo purificaciones exhaustivas de la enzima y evitando el uso de soportes que no estén disponibles en el mercado. A todo esto hay que añadir que cada enzima es diferente y que, por lo tanto, no existe una solución universal que nos valga para cualquier proceso. Teniendo en cuenta las necesidades actuales de la industria y la sociedad, la inversión en investigación en el desarrollo de sistemas de inmovilización sigue en auge, desarrollándose metodologías cada vez más sencillas y eficaces.



### 10.1. Caracterización de LSLt-BTL2 y LSLt-EGFP

Los estudios de inmovilización sobre partículas de agarosa se llevaron a cabo con dos proteínas distintas fusionadas a LSLt: (1) un mutante sin cisteínas de la lipasa termófila de *Geobacillus thermocatenolatus* (Godoy *et al.*, 2011) (LSLt-BTL2) y (2) la proteína fluorescente verde mejorada de *A. victoria* (LSLt-EGFP). Ambas proteínas se purificaron sin problema con el método descrito previamente, obteniéndose rendimientos de producción de cerca de 10 mg/L de cultivo.

La caracterización de la actividad hidrolítica de LSLt-BTL2 reveló que la etiqueta de fusión no afecta a las propiedades catalíticas de la enzima, al igual que ocurre en LSLt-Bgl, indicando que BTL2 está bien plegada y que LSLt no interfiere en el mecanismo de catálisis (Tabla 7.3).

Enzima	$V_{\text{máx}}$ (U/mg)	$K_m$ (mM)	$K_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{cat}}/K_m$ ( $\text{s}^{-1} \text{M}^{-1}$ ) $\times 10^4$
BTL2 <sup>a</sup>	95,9 $\pm$ 0,2	1,46 $\pm$ 0,51	69,5 $\pm$ 0,2	4,7
LSLt-BTL2 <sup>b</sup>	78 $\pm$ 3	1,2 $\pm$ 0,1	56,8 $\pm$ 2,2	4,8
Ag(4BCL):LSLt-BTL2 <sup>c</sup>	75 $\pm$ 4	2,3 $\pm$ 0,1	54 $\pm$ 3	2,3

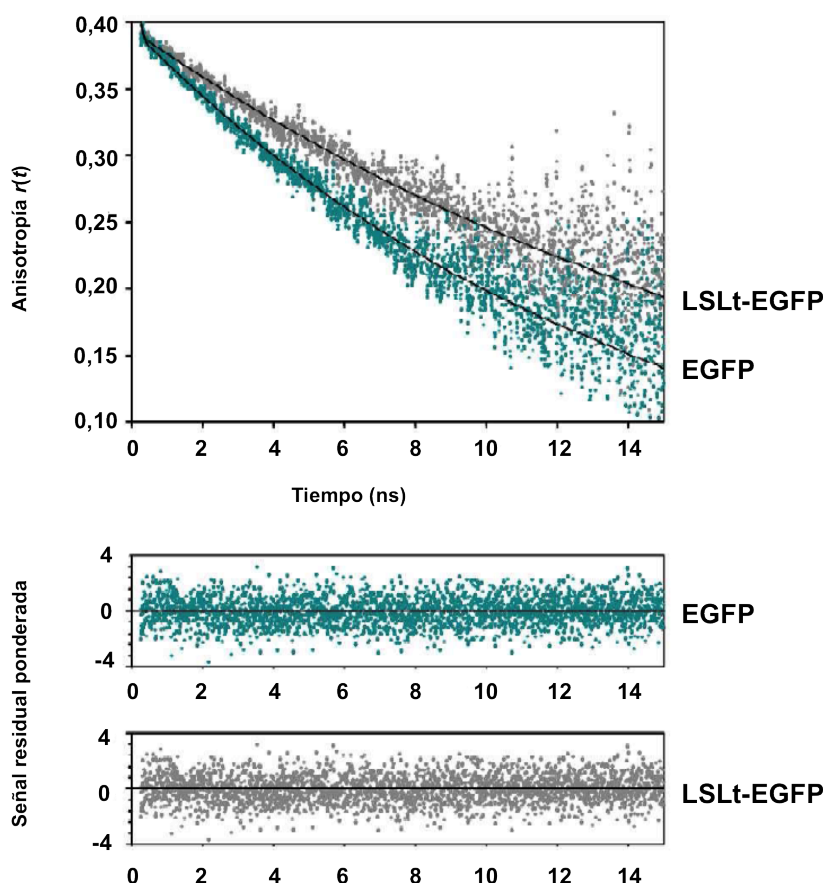
<sup>a</sup> Variante de BTL2 donde los dos residuos de Cys han sido sustituidos por residuos de Ser.

<sup>b</sup> Proteína de fusión LSLt-BTL2.

<sup>c</sup> Proteína de fusión inmovilizada sobre una matriz de agarosa de tipo 4BCL.

**Tabla 7.3.** Parámetros cinéticos de la actividad lipasa de diferentes formas de la enzima BTL2.

Por otro lado, además de la caracterización mediante dicroísmo circular que se llevó a cabo con LSLt-EGFP, se estudió el comportamiento espectroscópico en solución de la proteína de fusión y de la versión sin etiquetar (EGFP). Para ello se midió, entre otros parámetros, la vida media del estado excitado de EGFP, obteniéndose en ambos casos un valor de 2,85 ns, lo que demuestra que LSLt-EGFP tiene el mismo comportamiento espectroscópico que EGFP (Uskova *et al.*, 2000). No obstante, tal y como se espera para una proteína de mayor masa molecular, la anisotropía de fluorescencia de LSLt-EGFP (45 kDa) decae más lentamente que la anisotropía de la EGFP libre (27 kDa) (Figura 7.18). Estos resultados, en su conjunto, indican que LSLt no afecta al comportamiento espectroscópico intrínseco de EGFP, volviéndose a demostrar que LSLt no interfiere de ninguna manera en el plegamiento del módulo C-terminal.

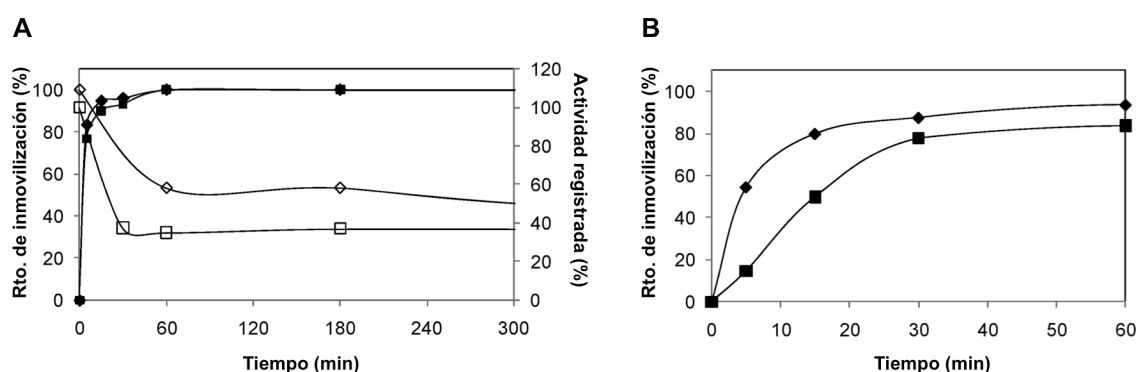


**Figura 7.18.** Decaimiento de la anisotropía de fluorescencia de EGFP y LSLt-EGFP. Los experimentos se llevaron a cabo a una concentración de 500 nM en tampón TSA y a 22 °C. Las longitudes de onda empleadas fueron de 482 nm para la excitación ( $\lambda_{ex}$ ) y 515 nm para la emisión ( $\lambda_{em}$ ). El ajuste de los datos se representa mediante una *línea continua*.

## 10.2. Caracterización de la inmovilización de proteínas recombinantes a partículas de agarosa

Las proteínas de fusión LSLt-BTL2 y LSLt-EGFP fueron inmovilizadas en dos tipos de matrices de agarosa distintas: agarosa 4B y agarosa 4BCL. El primer tipo es simplemente un polímero lineal de agarosa; mientras que el segundo, sin embargo, hace referencia a una agarosa sobre la que se ha producido una reacción de entrecruzamiento con epíclorhidrina con el fin de aumentar la estabilidad de las partículas bajo determinadas condiciones. Cabe señalar también que ni BTL2 ni EGFP tienen capacidad para unirse a ninguno de los dos tipos de agarosa, por lo que la inmovilización de las proteínas de fusión tiene siempre lugar a través de los sitios de unión del módulo lectina. En la **Figura 7.19** se muestran los parámetros de la inmovilización de LSLt-BTL2. Cuando se trabaja a baja carga (30 U/g de soporte), la proteína de fusión se une cualitativamente igual a ambos tipos de agarosa (**Figura 7.19A**). Debido a que las partículas de agarosa son capaces de unir más cantidad de proteína (hasta un máximo de 43 mg de LSLt-BTL2 por gramo de soporte seco), se repitieron los mismos ensayos a alta carga (400 U/g de soporte). En este caso, sin embargo, la inmovi-

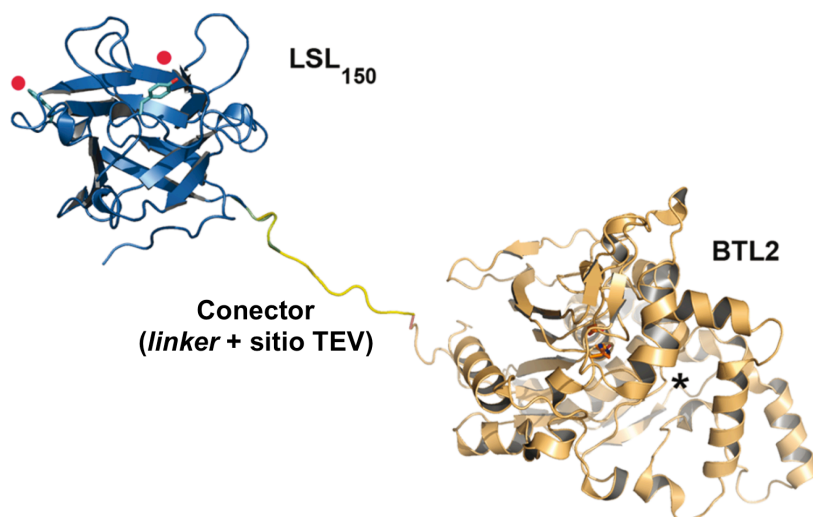
lización se vio ligeramente favorecida en el caso de la agarosa 4B (**Figura 7.19B**). Las diferentes cinéticas de inmovilización observadas probablemente se deben a las diferencias en la estructura química de cada tipo de agarosa. El tratamiento de entrecruzamiento de la agarosa 4BCL altera la estructura química de las unidades de D-galactosa y 3,6-anhidro-L-galactosa, afectándose, en consecuencia, la interacción establecida con la lectina (Lopez-Gallego *et al.*, 2012).



**Figura 7.19.** Inmovilización de LSLt-BTL2 a diferentes tipos de matrices de agarosa. LSLt-BTL2 fue inmovilizada en agarosa 4B (rombos) y agarosa 4BCL (cuadrados). **A.** El rendimiento de la inmovilización (símbolos rellenos) y la actividad registrada (símbolos abiertos) fueron calculados para cada experimento de inmovilización. La inmovilización se llevó a cabo incubando 30 U de actividad de LSLt-BTL2 por gramo de soporte en un ratio de inmovilización de 0,25 m<sup>2</sup> de soporte por mL de mezcla de inmovilización. **B.** La inmovilización se llevó a cabo incubando 400 U de actividad de LSLt-BTL2 por gramo de soporte en un ratio de inmovilización de 0,25 m<sup>2</sup> de soporte por mL de mezcla de inmovilización. En ambos casos, la inmovilización tiene lugar a pH 7,0 y 25 °C. Los patrones de inmovilización fueron bastante similares entre ambos tipos de matrices de agarosa, y solo cuando las condiciones de inmovilización se forzaron para inmovilizar grandes cantidades de enzima, una inmovilización relativamente más rápida se observó en la agarosa 4B. Los datos que se muestran se han obtenido a partir de la media de triplicados cuyo error no superó el 5%.

Sorprendentemente, la actividad enzimática registrada de la lipasa inmovilizada disminuyó como resultado del proceso de inmovilización. Las actividades medidas fueron de aproximadamente un 60 y un 50% de la actividad total de LSLt-BTL2 en solución cuando la proteína se inmovilizó en agarosa 4B y agarosa 4BCL, respectivamente (**Figura 7.19A**). Normalmente, una reducción en la actividad enzimática a causa de la inmovilización puede estar relacionada con dos factores diferentes: (a) inactivación de la enzima debido al protocolo de inmovilización o (b) problemas de difusión del sustrato entre los poros del soporte. En este caso, la primera opción debería descartarse, puesto que *a priori* la orientación de la proteína LSLt-BTL2 es favorable y evita que el sitio activo de BTL2 interaccione con el soporte directamente, ya que el sitio por donde entra el sustrato queda orientado en el lado opuesto al de los sitios de reconocimiento de azúcares del módulo LSLt (**Figura 7.20**). Parece más plausible, por tanto, pensar que la reducción en la actividad se debe a problemas de accesibilidad del sustrato. De hecho, el sustrato utilizado en los ensayos (pNPB) se utilizó en condiciones de no saturación (500 μM) debido a su baja solubili-

dad en soluciones acuosas. Asimismo, se comprobó que si la proteína inmovilizada se disociaba del soporte con lactosa, la actividad enzimática se recuperaba por completo, lo que justifica los problemas de difusión del sustrato pNPB. En este sentido, el valor aparente de la constante  $K_m$  obtenido para la proteína LSLt-BTL2 inmovilizada es prácticamente el doble que el determinado para la misma proteína en solución (**Tabla 7.3**), lo que sugiere nuevamente que durante la inmovilización existen mayores restricciones en la difusión del sustrato.



**Figura 7.20.** Modelo atómico supuesto de la estructura de LSLt-BTL2. Este modelo se construyó teniendo en cuenta las estructuras cristalográficas de LSL<sub>150</sub> y BTL2 (PDB 2W22) con una región conectora que contiene 11 aminoácidos que provee flexibilidad a la zona entre la lectina y la enzima. Los dos círculos rojos indican los dos sitios operativos de unión de azúcares y el asterisco, el sitio activo de entrada del sustrato (los residuos que participan en la catálisis están representadas mediante el modelo de varillas).

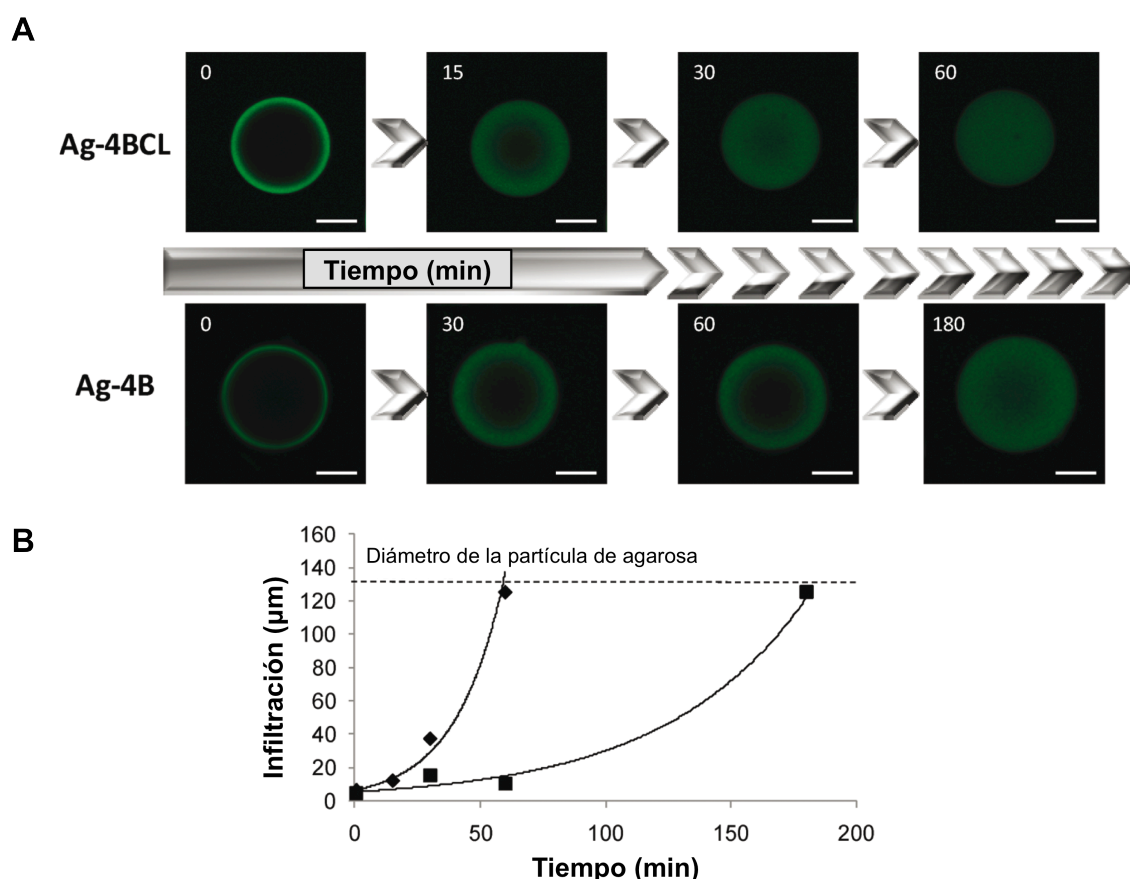
El estudio de la distribución de la proteína de fusión en las partículas de agarosa podría explicar la pérdida aparente de actividad a pesar de que la enzima no se inactiva por la mera inmovilización (Ladero *et al.*, 2001). En este sentido, se ha descrito que la distribución heterogénea de enzimas promueve una mayor eficacia en la catálisis en comparación con aquellos casos en los que las enzimas se distribuyen homogéneamente a través de las partículas y los poros del soporte (Borchert y Buchholz, 1984). Así, en los sistemas que muestran una distribución heterogénea, la eficacia catalítica recae en la mejor capacidad de difusión de los reactantes hacia el centro catalítico. En el sistema de inmovilización objeto de estudio, la distribución de la enzima depende única y exclusivamente del modo de interacción del módulo lectina con la agarosa, por lo que su estudio es clave en la comprensión de la relación entre la eficacia de la catálisis, la distribución de la enzima y la difusión del ligando. En este sentido, se utilizó la proteína de fusión LSLt-EGFP y los dos tipos de agarosa para llevar a cabo ensayos de microscopía confocal láser de barrido (CLSM). Los resultados demostraron que la proteína de fusión, aunque se une rápidamente a ambos tipos de agarosa, de un modo similar a una inmovilización covalente, acaba adoptando una distribu-

ción homogénea. En principio, este hecho no obedece a lo observado con otras proteínas, que al unirse rápidamente a los soportes, muestran una distribución heterogénea, estando generalmente localizadas en la superficie de las partículas (Bolivar *et al.*, 2011).

Con la intención de comprender el mecanismo de distribución de las proteínas de fusión en los soportes de agarosa, la unión se estudió de forma dinámica. En primer lugar, LSLt-EGFP fue rápidamente inmovilizada en ambos tipos de agarosa a 4 °C. Después de un minuto, la suspensión con la proteína inmovilizada se filtró y se lavó para eliminar la proteína no unida. Por último, las partículas de agarosa con LSLt-EGFP inmovilizada se diluyeron en una solución acuosa y su fluorescencia fue registrada a diferentes intervalos de tiempo. Inesperadamente, se observó cómo las proteínas migraron hacia el interior de las partículas de agarosa desde la superficie de éstas durante el periodo de postinmovilización (**Figura 7.21**). De este modo, mientras que a tiempos cortos LSLt-EGFP se distribuye heterogéneamente, apareciendo solamente en la superficie exterior de las partículas, a tiempos más largos, LSLt-EGFP adquiere una distribución más homogénea, migrando a través de los poros hasta cubrir la superficie interna de las partículas de agarosa. El comportamiento dinámico de la distribución de LSLt-EGFP revela un modo de unión entre LSLt y la agarosa basado en un proceso reversible y rápido de asociación-disociación. La reversibilidad de este proceso es la que permitiría la migración observada desde una distribución heterogénea en la superficie de las partículas hacia una distribución homogénea. La migración desde las zonas más pobladas de la agarosa hacia las más desocupadas se debería principalmente a la presencia de unidades de galactosa repartidas a lo largo de la superficie de todo el soporte. En este sentido, al disociarse las proteínas de fusión, éstas tenderían a interactuar con las zonas menos pobladas, que coinciden con aquellas donde la densidad de unidades de galactosa libres es más alta. El equilibrio final se alcanzaría cuando toda la agarosa estuviera igual de “ocupada”. Cabe señalar, que la adquisición de una distribución homogénea se consigue más rápidamente en la agarosa 4BCL que en la 4B. Probablemente, este comportamiento vuelva a deberse a la diferencia en la estructura química de ambas agarosas. Teniendo en cuenta que la unión a la agarosa 4BCL era relativamente más débil, probablemente la disociación de ésta también esté más favorecida y, por tanto, se alcanza más rápidamente la homogeneidad.

Esta es la primera vez que la distribución homogénea de una proteína inmovilizada se consigue sin necesidad de trabajar saturando el sistema, es decir, trabajando a baja carga. Esto es posible gracias al mecanismo de migración que experimenta LSLt durante el periodo de postinmovilización. Esta característica tiene potenciales aplicaciones, sobre

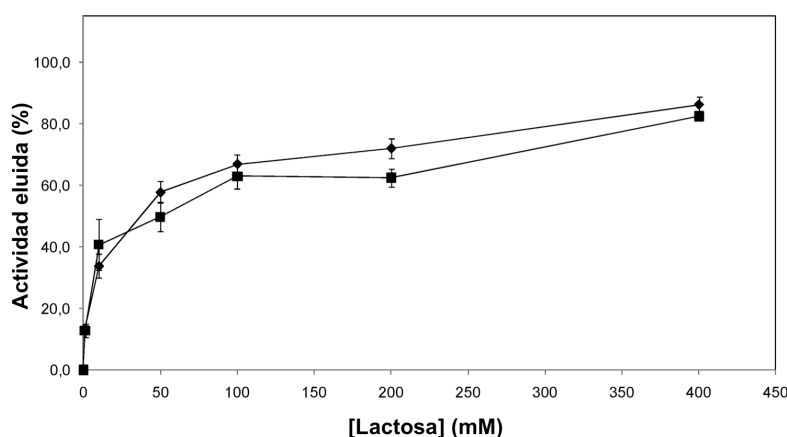
todo en el desarrollo de chips para detectar analitos importantes en medicina o química medioambiental (Jokerst *et al.*, 2011) o para utilizar en combinación con otros métodos de inmovilización que promueven una distribución heterogénea y así conseguir complejos ordenados de biomoléculas, como, por ejemplo, complejos multienzimáticos (Logan *et al.*, 2007; Teeuwen *et al.*, 2009).



**Figura 7.21.** Dependencia con el tiempo de la distribución espacial de LSLt-EGFP sobre las partículas de los diferentes tipos de matrices de agarosa. La inmovilización se llevó a cabo mediante la incubación de 1 mg de LSLt-EGFP por gramo de cada tipo de soporte. Cuando LSLt-EGFP se unió cuantitativamente, la agarosa se filtró a vacío y se resuspendió en 25 mM fosfato de sodio pH 7,0 y posteriormente se analizó con CLSM. **A.** Las imágenes de CLSM se grabaron a diferentes intervalos de tiempo. **B.** La infiltración de LSLt-EGFP fue calculada a diferentes tiempos, midiendo el tamaño del radio ocupado por el 90% del total de la fluorescencia dada en cada caso para la proteína unida a agarosa 4B (*cuadrados*) o agarosa 4BCL (*rombos*). Los datos experimentales se ajustaron a la ecuación matemática  $D_i = A \cdot e^{V_d \cdot t}$  para determinar la distribución de velocidad ( $V_d$ ). Se puede observar que en ambos tipos de agarosa la proteína LSLt-EGFP se infiltró cada vez más en los poros de la superficie de la partícula. La distribución homogénea completa llevó más tiempo en el caso de la inmovilización a agarosa 4B. El tamaño de las partículas de agarosa representadas en la figura ronda los 125  $\mu\text{m}$ , siendo la escala (*barra blanca*) de 50  $\mu\text{m}$ .

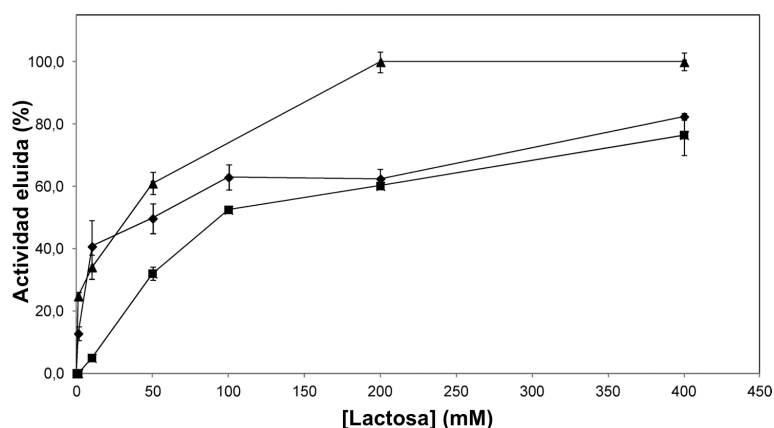
### 10.3. Efecto de diferentes variables en la unión de LSLt a agarosa

Considerando la capacidad que tiene LSLt para disociarse de la agarosa en presencia de lactosa, se estudió la dependencia de la unión con respecto a la concentración del disacárido para evaluar así la afinidad aparente que LSLt muestra hacia las partículas de agarosa. En la **Figura 7.22** se muestra que LSLt-BTL2 tiene un perfil de elución muy parecido para ambos tipos de agarosa 4B y 4BCL, lo que sugiere que LSLt se une con la misma fuerza a las matrices, al menos, desde un punto de vista macroscópico. Teniendo en cuenta que la agarosa con entrecruzamiento presenta mejores propiedades mecánicas, este resultado podría abrir un nuevo abanico de posibilidades para aplicar LSLt en la biocatálisis aplicada, como un método genérico de inmovilizar proteínas.



**Figura 7.22.** Capacidad de unión del módulo LSLt a diferentes tipos de matrices de agarosa. La proteína LSLt-BTL2, inmovilizada a 4 °C durante 15 min (100% de inmovilización), se eluyó a diferentes concentraciones de lactosa (0-400 mM). Este experimento permitió evaluar la afinidad aparente del módulo LSLt a las unidades de galactosa que forman la estructura de la agarosa. Las dos matrices estudiadas fueron agarosa 4B (*rombos*) y agarosa 4BCL (*cuadrados*). Los datos se muestran con las barras de error ( $\sigma$ ), determinadas a partir de experimentos triplicados.

Asimismo, también se analizó el efecto de la temperatura y el tiempo a la fuerza de unión de LSLt a la agarosa. Con este propósito, LSLt-BTL2 fue inmovilizada en agarosa 4BCL durante distintos periodos de tiempo (15 min y 18 h) a dos temperaturas (4 y 25 °C) para posteriormente llevar a cabo la elución con diferentes concentraciones de lactosa. De esta manera, se vio que la temperatura es uno de los factores que más afectan a la fuerza de unión (**Figura 7.23**). Por ejemplo, el 60% de la proteína inmovilizada a 4 °C durante 18 h eluyó con 50 mM lactosa, mientras que fue necesario 200 mM lactosa para registrar el mismo porcentaje de actividad cuando LSLt-BTL2 se inmovilizó a 25 °C durante el mismo tiempo. Aunque el tiempo de incubación no afecta tanto como la temperatura a la fuerza de la unión, es interesante resaltar cómo se necesita menor concentración de lactosa en la elución cuando el tiempo de inmovilización ha sido más corto.



**Figura 7.23.** Efecto de la temperatura y el tiempo en la capacidad de unión del módulo LSLt a la agarosa 4BCL. La elución de la proteína de fusión LSLt-BTL2 se llevó a cabo usando diferentes concentraciones de lactosa (0-400 mM). Este experimento permite evaluar cómo de fuerte se une el módulo LSLt a las unidades de galactosa que forman la estructura de la agarosa 4BCL. La inmovilización tuvo lugar a diferentes temperaturas, 25 °C (*cuadrados*) y 4 °C (*triángulos*), durante 18 horas, observándose la mayor unión cuando la unión se llevó a cabo a la mayor temperatura. LSLt-BTL2 también fue inmovilizado a 25 °C durante 15 min (*rombos*), observándose menor unión en tiempos más cortos. Los datos se muestran con las barras de error ( $\sigma$ ), determinadas a partir de experimentos triplicados.

Los resultados anteriores acerca del efecto de la temperatura permiten jugar con este parámetro en función del objetivo con el que se produzca la proteína de fusión. Cuando el propósito es purificar una proteína interesa que la unión a la matriz de agarosa no sea muy intensa ya que la proteína debe eluirse posteriormente. Sin embargo, cuando se quiere inmovilizar una proteína para llevar a cabo un proceso específico, la unión ha de ser fuerte con el fin de evitar al máximo el *leaching*. Teniendo esto en cuenta, convendría usar bajas temperaturas para la purificación y temperaturas más elevadas en la inmovilización.

#### 10.4. Estabilidad del módulo LSLt bajo diferentes condiciones

Como se ha visto a lo largo de este trabajo, la unión de LSLt a la agarosa depende en último término de la estructura tipo trébol  $\beta$  y, en particular, a la disposición espacial de los residuos de los dos sitios operativos de la lectina en el estado nativo ( $\beta$  y  $\gamma$ ). Por este motivo, la pérdida total o parcial de la estructura del módulo LSLt acarrearía la disociación de la proteína de fusión de la matriz o soporte. En la biocatálisis industrial, las condiciones de reacción pueden llegar a ser drásticas y afectar a la estructura de las proteínas. Con la finalidad de evaluar la estabilidad de la estructura del módulo lectina LSLt, LSLt-BTL2 se inmovilizó de forma óptima en agarosa 4BCL y se incubó bajo diferentes condiciones de pH y temperatura y en presencia de solventes orgánicos. Es importante considerar en este ensayo, que la proteína LSLt-BTL2 en solución retuvo el 100% de la actividad enzimática, por lo que los resultados obtenidos dependen exclusivamente de la estabilidad del dominio LSLt.



A la vista de la **Tabla 7.4** se observa que el módulo LSLt es muy estable a altas temperaturas, como ya se vio en los perfiles de desnaturalización térmica, y que tolera relativamente bien altas concentraciones de acetonitrilo (más de un 90% de la proteína inmovilizada permanece unida a las partículas después de una incubación de hasta 60 °C o hasta un 60% de acetonitrilo). En cambio, la unión a la agarosa se vio fuertemente afectada en condiciones de pH básico, donde eluye el 100% de la proteína, o en presencia de surfactantes no iónicos como el Triton X-100. Por tanto, las condiciones alcalinas y los detergentes parecen promover cambios en el plegamiento tipo trébol  $\beta$  que conducen a la disociación del soporte. Aun así, esta metodología de inmovilización es bastante robusta para una gran variedad de condiciones de reacción. El método de inmovilización basado en el uso de LSLt puede ser usado en reacciones de química fina donde los sustratos frecuentemente presentan problemas de solubilidad y los solventes orgánicos son requeridos en el medio de reacción. Más allá, las enzimas inmovilizadas vía LSLt podrían también emplearse en ambientes de reacción que transcurriesen a una temperatura media-alta (40–60 °C), como, por ejemplo, en la industria alimentaria donde es necesario elevar la temperatura para evitar las contaminaciones microbianas. Por estos motivos, la robustez de la unión LSLt-agarosa, la simplicidad del proceso de inmovilización y el relativo bajo coste que presenta el soporte no activado de agarosa hacen de esta metodología una buena alternativa para inmovilizar biocatalizadores de una forma barata y efectiva con potenciales aplicaciones en la industria actual.

Condiciones de elución	Actividad de BTL2 eluida (%)*
Fosfato de sodio pH 7,0 y 25 °C	0
Fosfato de sodio pH 7,0 y 45 °C	7
Fosfato de sodio pH 7,0 y 60 °C	3
Fosfato de sodio pH 5,0 y 25 °C	20
Fosfato de sodio pH 10,0 y 25 °C	100
Fosfato de sodio pH 7,0 y 25 °C en presencia de 0,2% de Triton X-100	60
20% acetonitrilo pH 7,0 y 25 °C	10
40% acetonitrilo pH 7,0 y 25 °C	0
60% acetonitrilo pH 7,0 y 25 °C	5

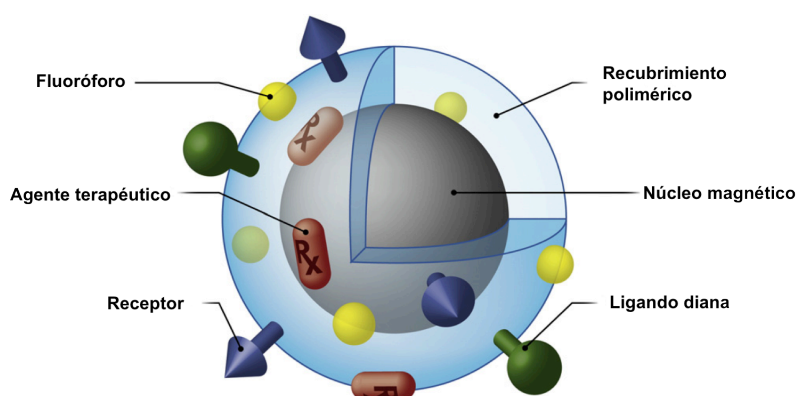
\*La actividad de la fracción eluida fue medida por triplicado, no siendo el error nunca superior al 10%.

**Tabla 7.4.** Elución de LSLt-BTL2 de la agarosa 4BCL bajo diferentes condiciones.

## 11. Inmovilización en nanopartículas magnéticas

La nanotecnología es considerada por muchos como la nueva revolución científica del momento. Un ejemplo de esta revolución es el descubrimiento de las nanopartículas magnéticas (NPM), que han permitido el desarrollo de nuevas tecnologías a nanoescala. Se han encontrado para las nanopartículas aplicaciones en diferentes campos, tales como la electrónica, las comunicaciones, la óptica, la química energética y, por supuesto, la biología (Moros *et al.*, 2010). Por ejemplo, actualmente hay dos líneas de investigación que abogan por el uso de las NPM como la próxima generación de agentes de contraste para la toma de imágenes por resonancia magnética (IRM) y como transportadores de fármacos sitio-específicos en terapias para ciertas enfermedades (Pankhurst *et al.*, 2003; Dobson, 2006). Con un amplio rango de aplicaciones en la detección, diagnóstico y tratamiento de dolencias, tales como el cáncer, enfermedades cardiovasculares y enfermedades neurológicas, las NPM parecen jugar un papel fundamental en los tratamientos del futuro (Sun *et al.*, 2008).

Las NPM son, por lo general, cristales esféricos con un tamaño de 10-20 nm y frecuentemente están elaboradas con un núcleo de óxido de hierro, aunque hay NPM fabricadas con amplia variedad de materiales. En función de la aplicación, las NPM se recubren con algún material que no modifique sus propiedades magnéticas, pero que aumente su solubilidad en soluciones acuosas y prepare la superficie para cualquier tipo de marcaje (**Figura 7.24**). Por ejemplo, las NPM con aplicaciones en biomedicina deben disponer de recubrimientos biocompatibles que se mantengan estables bajo condiciones fisiológicas. Este diseño modular (núcleo + recubrimiento) permite diseñar gran variedad de tipos de NPM con funcionalidades varias.



**Figura 7.24.** Nanopartícula magnética que exhibe diversos ligandos que posibilitan su aplicación en diferentes campos.

Las NPM representan una opción atractiva no solo en biomedicina, sino también en campos, como la biotecnología. De hecho, suponen un sistema ideal para la inmovilización de enzimas (Rossi *et al.*, 2004). Como se ha visto en el apartado anterior, la inmovilización

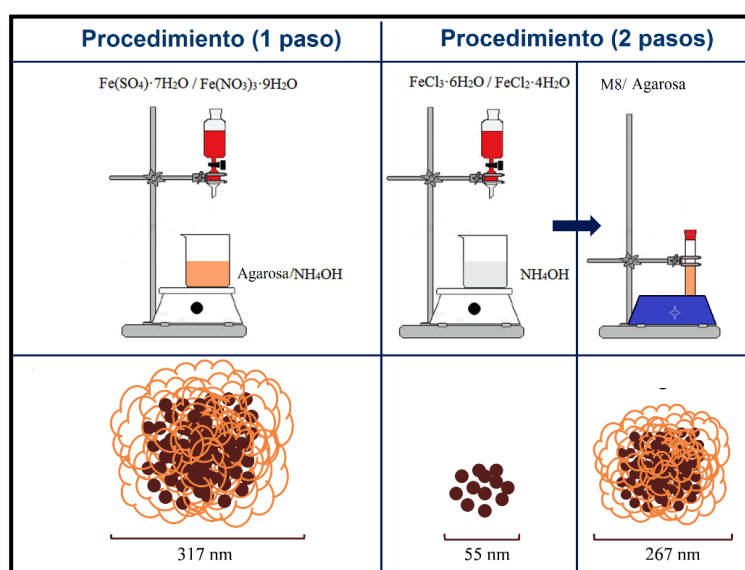
de enzimas incrementa la estabilidad de éstas y facilita su reutilización. Asimismo, el gran área expuesta en relación con el volumen se traduce en una reducción de la transferencia de masa, así como la existencia de menos limitaciones de difusión de los sustratos para una carga máxima de las NPM (Santana *et al.*, 2012). Por otro lado, las características superparamagnéticas de las NPM facilitan enormemente la separación de las fases sólida y líquida, por lo que al producto no le acompañan contaminantes procedentes de la reacción. No hay que olvidar que la variedad de recubrimientos hacen que éstos puedan ser fácilmente modificados para añadir grupos funcionales de diferente naturaleza y así poder llevar a cabo la inmovilización con gran diversidad de métodos (Zhang *et al.*, 2011).

Las NPM se pueden convertir en sistemas muy versátiles para estudiar las interacciones lectina-carbohidrato. Se han desarrollado recubrimientos en los que se exponen diferentes azúcares (Marradi *et al.*, 2013) y en los que se inmovilizan lectinas con los sitios de unión expuestos para reconocer azúcares, actuando así como biosensores (Dai *et al.*, 2006). Hay que destacar que el recubrimiento con carbohidratos estabiliza las nanopartículas y evita problemas de agregación (Moros *et al.*, 2010). Y es que uno de los retos en la producción de nanopartículas es evitar la agregación de éstas, ya que si, por ejemplo, se quisieran utilizar *in vivo*, los agregados podrían traducirse en problemas serios tales como la trombosis. En este contexto, hemos desarrollado un sistema basado en el uso de NPM de óxido de hierro derivatizadas con un recubrimiento de agarosa con el fin de utilizarlas como soporte para la inmovilización de proteínas de fusión mediante la interacción lectina-agarosa. Es interesante destacar que en la última década, el número de nanopartículas magnéticas funcionalizadas con azúcares y aplicadas al desarrollo de aplicaciones con lectinas ha crecido a buen ritmo y se espera que se incremente en los próximos años.

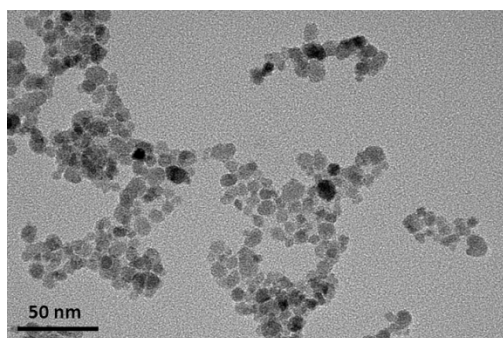
### 11.1. Las lectinas se unen a nanopartículas recubiertas de agarosa

En colaboración con el grupo de investigación de la Dra. M<sup>a</sup> del Puerto Morales Herrero del ICMN (CSIC), se sintetizaron nanopartículas magnéticas con un núcleo de maghemita (óxido de hierro) y recubiertas de agarosa (NPA) mediante dos métodos de coprecipitación distintos (Tartaj *et al.*, 2003) (**Figura 7.25**). El primer método consiste en la síntesis de las nanopartículas en un solo paso; mientras que en el segundo procedimiento (en dos pasos), las nanopartículas de maghemita se obtienen vía coprecipitación y son funcionalizadas posteriormente con la agarosa. De esta manera, se obtuvieron dos tipos de NPM con un diámetro de 8 nm y un recubrimiento de agarosa, denominadas NPA-1 y NPA-2, respectivamente (**Figura 7.26**). En ambos casos, se llevaron a cabo estudios fisicoquímicos para

determinar algunas propiedades, como el punto isoelectrico, el radio hidrodinámico y el potencial  $\zeta$ , que permitieron corroborar que las NPM estaban efectivamente recubiertas de agarosa. La agarosa es un polímero hidrofílico que presenta gran resistencia química en un amplio rango de pH (0–14) y que está siendo utilizado actualmente como material de soporte para diferentes tipos de experimentos. No obstante, el principal problema de las partículas de agarosa es su fragilidad y su susceptibilidad al daño por agitación (Safdarian *et al.*, 2013).



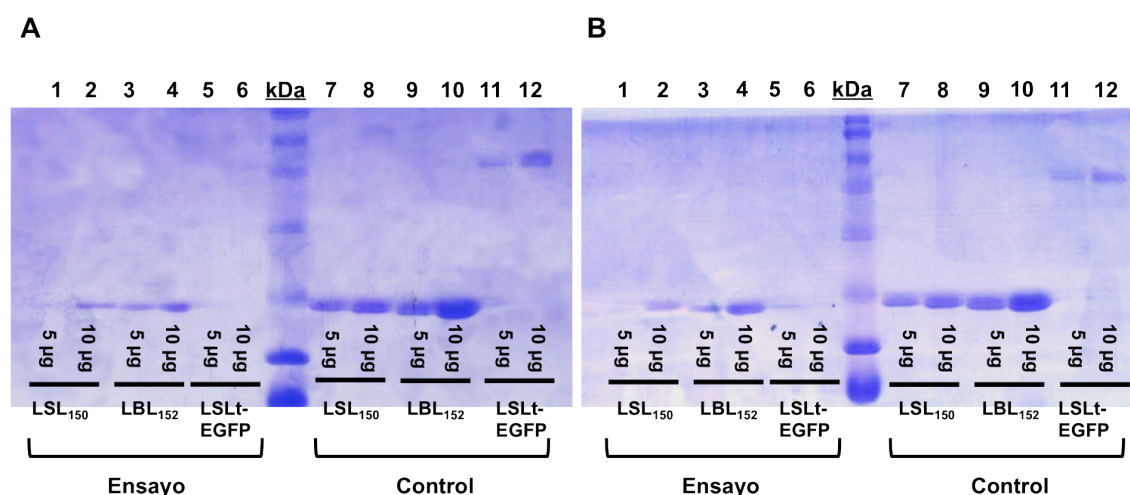
**Figura 7.25.** Representación esquemática de los procedimientos empleados para la obtención de nanopartículas recubiertas de agarosa.



**Figura 7.26.** Imagen de TEM (Microscopía electrónica de transmisión) de las nanopartículas recubiertas de agarosa (NPA).

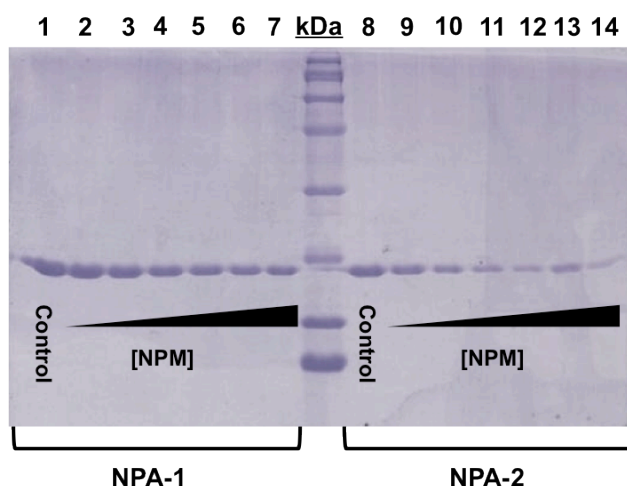
Uno de los mayores retos en el mundo de las nanopartículas es encontrar un modo eficaz para funcionalizar las NPM, es decir, anclar el ligando de interés y dotarlas de función. En la actualidad existen diferentes estrategias, desde uniones vía avidina-biotina hasta modificaciones covalentes para formar enlaces tipo amida, éster o disulfuro (Sun *et al.*, 2008). En este contexto, queremos desarrollar un nuevo sistema que aproveche la capacidad de reconocimiento de los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> para derivatizar NPM y otorgarles así una función determinada. Por esta razón, una vez se hubieron preparado las nanopartículas recubiertas de agarosa (NPA) se realizaron ensayos de unión con el fin de

comprobar si estas NPM permitían la inmovilización de los módulos lectina objeto de estudio en este trabajo. Los geles de electroforesis de la **Figura 7.27** muestran los resultados en los que se observa como la intensidad de las bandas disminuye con respecto a la de los controles cuando las proteínas (LSL<sub>150</sub>, LBL<sub>152</sub> y LSLt-EGFP) han sido incubadas previamente con las NPA. Esta diferencia es consecuencia directa de la adhesión de las lectinas a las NPA. Para una cantidad fija de NPA, se observa que en todos los casos las bandas del gel son más intensas en presencia de mayor cantidad de proteína, lo que sugiere que las NPA estén en condiciones de saturación. Curiosamente, con ambos tipos de NPA, la lectina LBL<sub>152</sub> se une peor que LSL<sub>150</sub> ya que las bandas correspondientes son más intensas. De hecho, en presencia de 5  $\mu$ g de lectina, LSL<sub>150</sub> está completamente unida, mientras que en el caso de LBL<sub>152</sub>, hay proteína en solución. Con respecto a LSLt-EGFP, la proteína se une completamente a las NPA.



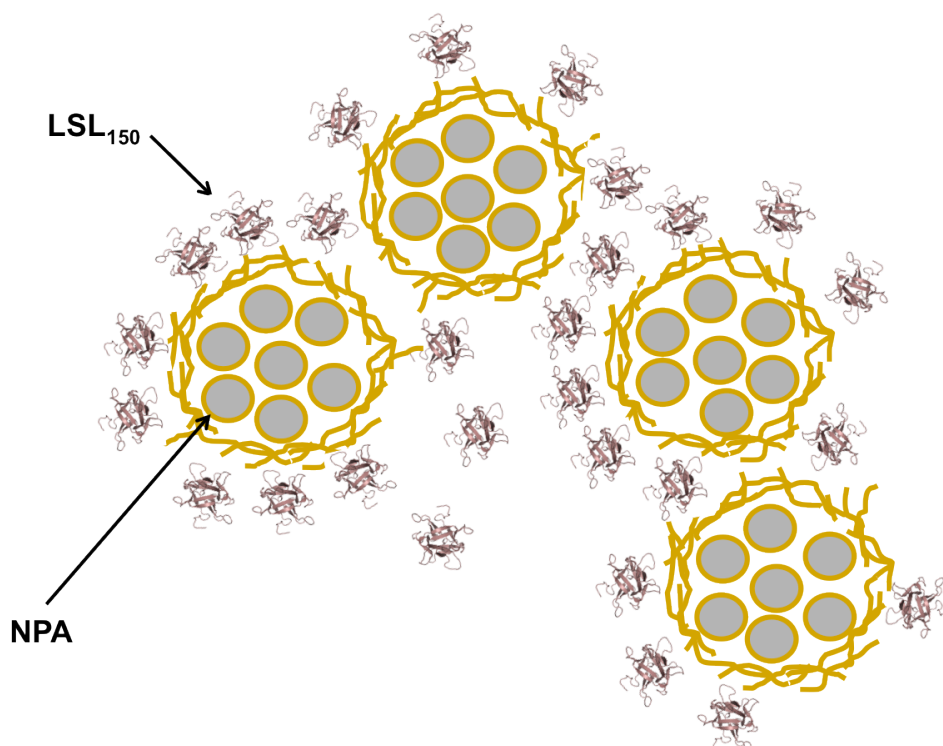
**Figura 7.27.** Ensayo de unión a dos tipos de NPM con recubrimiento de agarosa: NPA-2 (A) y NPA-1 (B). Tanto en el ensayo como en el control, las tres proteínas (LSL<sub>150</sub>, LBL<sub>152</sub> y LSLt-EGFP) fueron incubadas en presencia de las NPA durante 2 h a temperatura ambiente en un volumen final de 50  $\mu$ L en tampón TSA. En cada caso, se probaron dos cantidades distintas de proteína, 5 y 10  $\mu$ g totales, para un volumen fijo de NPA. En los geles de electroforesis, se muestra el análisis de la fracción soluble, que no contiene NPA.

Teniendo en cuenta que LSL<sub>150</sub> se une relativamente mejor a las NPA, se llevó a cabo una titulación con este módulo lectina y diferentes concentraciones de NPA para comprobar cuál de los dos tipos de NPA sintetizadas actúa como mejor sistema de inmovilización. El gel de la **Figura 7.28** demuestra que la interacción es dependiente de la cantidad de NPA en el medio, disminuyendo la intensidad de las bandas a medida que aumenta la concentración de NPA debido a la mayor interacción. Además, este resultado confirma que las NPA-2 se unen con mayor afinidad que las otras, puesto que la intensidad de las bandas es mucho menor.



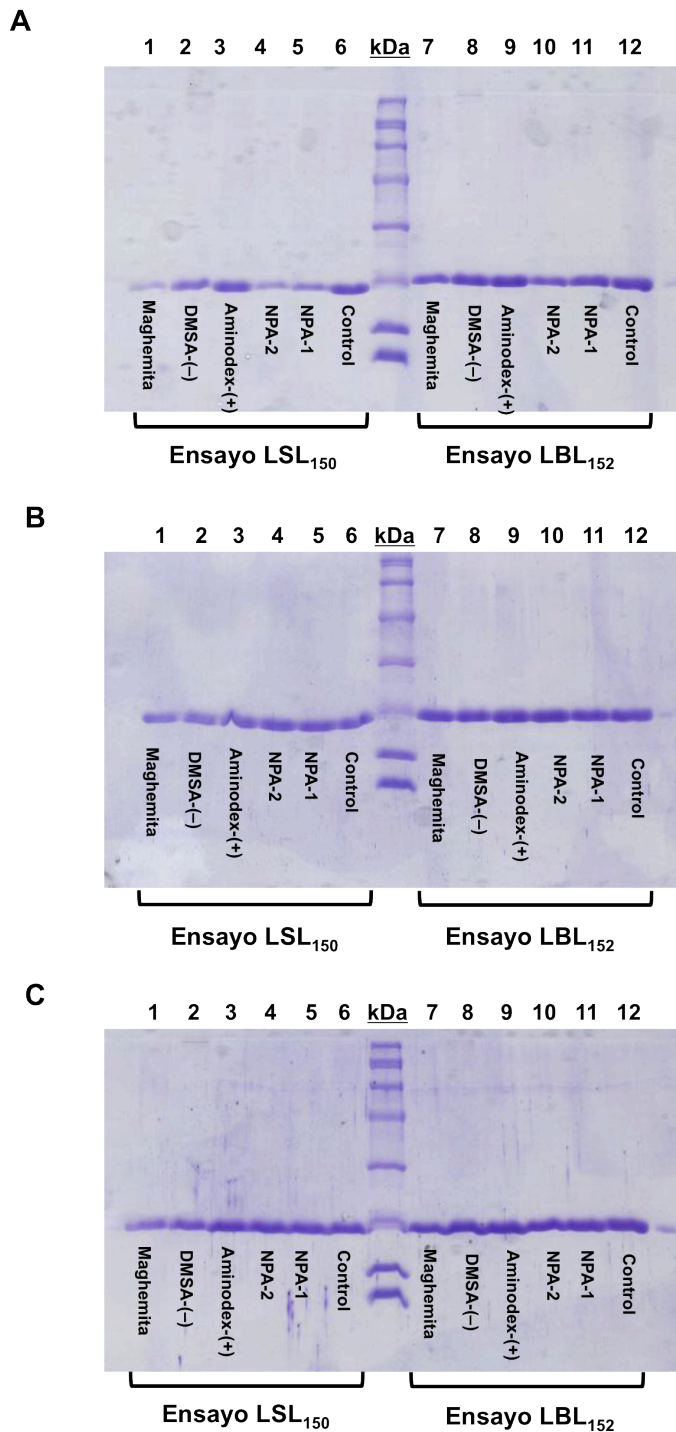
**Figura 7.28.** Dependencia de la unión de LSL<sub>150</sub> con la concentración de las NPM con recubrimiento de agarosa NPA-1 y NPA-2. La incubación se llevó a cabo con 10 µg de LSL<sub>150</sub> durante 2 h a temperatura ambiente en un volumen final de 70 µL en tampón TSA. En el gel de electroforesis, se muestra el análisis de la fracción soluble, que no contiene NPA.

Las NPA son equivalentes a las partículas de agarosa comerciales, puesto que la estructura química no se ve afectada durante el proceso de recubrimiento. Por esta razón, los resultados son consistentes con la capacidad de estas proteínas para reconocer motivos de galactosa. Curiosamente, la adición de las NPA a la solución de lectina produjo la inmediata precipitación de las NPA, yéndose éstas rápidamente hacia el fondo del tubo. Las NPA tienden de forma natural a precipitar por gravedad, pero más lentamente que en el comportamiento observado en presencia de las lectinas. Debido a que las lectinas tienen la capacidad de aglutinar eritrocitos, el comportamiento experimentado por las NPA podría deberse a este mismo fenómeno de aglutinación (**Figura 7.29**). Por ejemplo, este comportamiento ya ha sido descrito previamente para la lectina ConA, que en presencia de NPM recubiertas con glucosa (GNP) produce la aglutinación de éstas (Moros *et al.*, 2010).



**Figura 7.29.** Modelo esquemático donde se muestra el modo de unión de la lectina LSL<sub>150</sub> con las NPA a través de los sitios de unión. Algunas moléculas de LSL<sub>150</sub> podrían interactuar con más de una nanopartícula al mismo tiempo gracias a la multivalencia, provocando la aglutinación de las NPA.



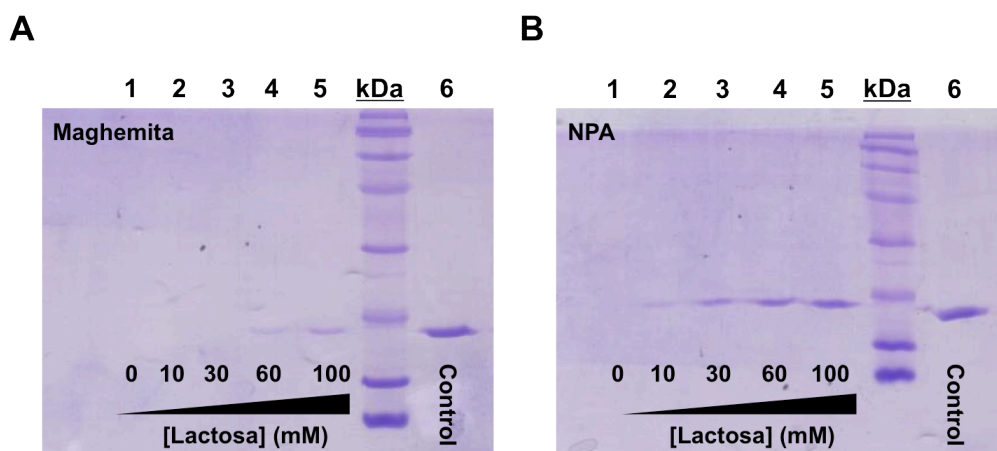


**Figura 7.30.** Ensayo de unión a un set de diferentes tipos de NPM. **A.** Incubación de 10  $\mu$ g de lectina durante 2 h a temperatura ambiente en un volumen final de 50  $\mu$ L en tampón TSA. **B.** Incubación de 10  $\mu$ g de lectina durante 2 h a temperatura ambiente en un volumen final de 50  $\mu$ L en tampón TSAL. **C.** Incubación de 10  $\mu$ g de lectina durante 2 h a temperatura ambiente en un volumen final de 50  $\mu$ L en tampón TSA seguida de una incubación con 0,1 M lactosa durante 1 h a temperatura ambiente. En los geles de electroforesis, se muestra el análisis de la fracción soluble, que no contiene NPM.

Para comprobar que la unión de las lectinas a las NPA tiene lugar de forma específica a través de los sitios de unión, se llevó a cabo un experimento, empleando diferentes tipos de NPM (**Figura 7.30**). En primer lugar (**Figura 7.30A**), a la vista del gel, se puede observar cómo LSL<sub>150</sub> se une a los dos tipos de NPA, mientras que no hay interacción con ninguna de las NPM cargadas. Sin embargo, curiosamente, se aprecia una fuerte interacción con las NPM de maghemita, que no poseen ningún tipo de recubrimiento. El comporta-

miento de LBL<sub>152</sub> es similar, pero la unión es mucho más baja, puesto que la intensidad de las bandas es mayor. Si este experimento se repite en presencia de lactosa desde el principio (**Figura 7.30B**) o se añade lactosa *a posteriori* (**Figura 7.30C**), la unión con las NPA se revierte por completo, lo que indica claramente que la unión transcurre específicamente a través de los sitio de unión de azúcares.

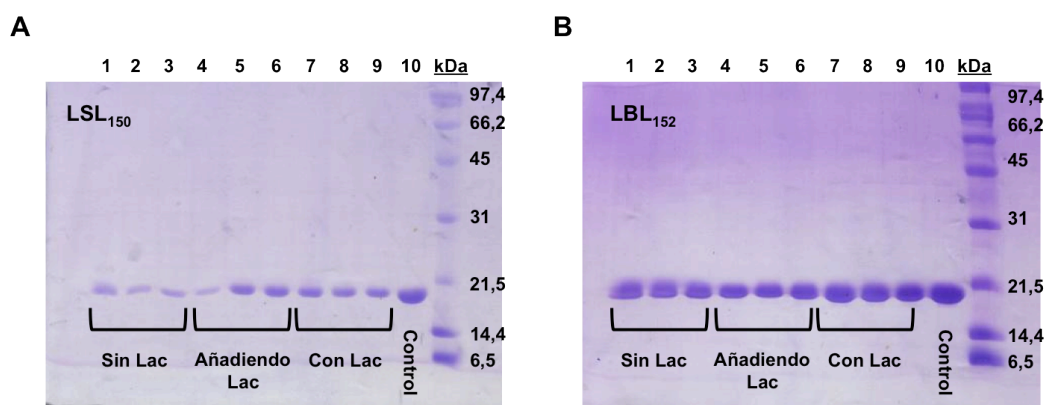
Tal y como se ha visto en los resultados anteriores, las lectinas interaccionaron con el núcleo de maghemita de las NPM. Con el fin de comprender cómo tiene lugar este tipo de unión, se hicieron dos titulaciones con lactosa para comparar la unión de LSL<sub>150</sub> (se usó esta lectina porque es la que mayor unión presenta) con los núcleos de maghemita y con las NPA (desde este momento se emplearon siempre las NPA-2 porque son las que mejor capacidad de inmovilización presentan). El resultado (**Figura 7.31**) demuestra que, mientras que LSL<sub>150</sub> reconoce a las NPA mediante los sitios específicos de unión de azúcares, la interacción con las NPM sin recubrimiento es inespecífica y no tiene lugar solamente vía los sitios de unión ya que no se consigue revertir del todo la unión en presencia de la lactosa. Por este motivo, aunque aparentemente los núcleos de maghemita son capaces de unir más cantidad de proteína, su uso como sistema para inmovilizar proteínas está limitado, puesto que no se controla la orientación de éstas en la superficie. Por otro lado, al igual que ocurría con las partículas de agarosa 4B y 4BCL, se observa que con 100 mM lactosa no se consigue recuperar el 100% de la lectina inmovilizada, pues la intensidad la banda de electroforesis con respecto al control es menor. Probablemente, si LSL<sub>150</sub> está implicada en la aglutinación de las NPA, la accesibilidad del ligando puede ser menor y, por tanto, es más difícil revertir la unión por completo.



**Figura 7.31.** Titulación con lactosa de la unión de la lectina LSL<sub>150</sub> a NPM sin recubrimiento (A) y recubiertas de agarosa (NPA) (B). La incubación se llevó a cabo durante 2 h a temperatura ambiente con una relación 1:10 de LSL<sub>150</sub>:NPM. La posterior incubación con 0,1 M lactosa se realizó durante 1 h a 4 °C. En los geles de electroforesis, se muestra el análisis de la fracción soluble, que no contiene NPM.



También quisimos comprobar si la unión de la lectina a la maghemita era estable con el tiempo. En este caso, el ensayo se realizó con los dos módulos lectina y en presencia o ausencia de lactosa. A la vista de los geles (**Figura 7.32**), se observa que LSL<sub>150</sub> queda retenido a las NPM sin recubrimiento y que esta unión no se revierte con el tiempo. Sin embargo, en presencia de lactosa, ya sea incubando desde el principio o añadiéndola *a posteriori*, la fracción de LSL<sub>150</sub> unida a las NPM es menor. Esto indicaría que la unión a la maghemita es inespecífica, es decir, no está orientada y, por lo tanto, la fracción que se libera con la lactosa se corresponde con aquella que queda orientada con los sitios de unión de azúcar de cara a las NPM. Curiosamente, a pesar de que LBL<sub>152</sub> es un homólogo estructural muy próximo, éste módulo lectina no experimenta ninguna interacción con las NPM sin recubrimiento. Esto valida de algún modo que la interacción lectina-NPM no se está llevando a cabo a través de los sitios de unión a azúcar y que con bastante probabilidad son otro tipo de interacciones las que entran en juego como, por ejemplo, interacciones electrostáticas.

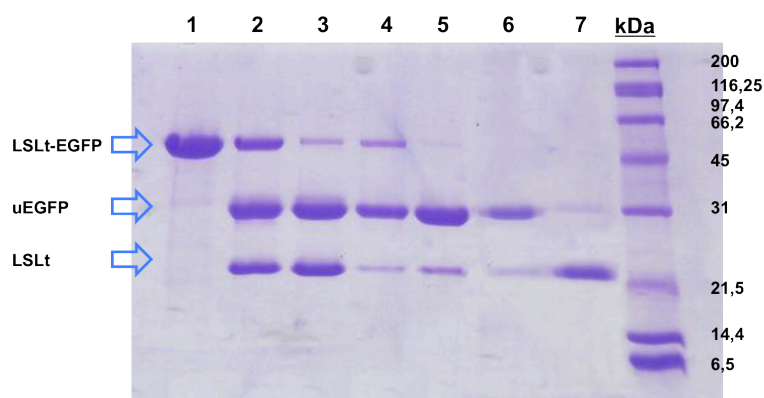


**Figura 7.32.** Estudio de la estabilidad de la unión de las lectinas LSL<sub>150</sub> (**A**) y LBL<sub>152</sub> (**B**) a las NPM sin recubrimiento (maghemita). El ensayo se hizo en tres condiciones diferentes: en ausencia total de lactosa (carriles 1, 2 y 3), añadiendo lactosa después de 2 horas de incubación a temperatura ambiente (carriles 4, 5 y 6) y en presencia de lactosa desde el principio (carriles 7, 8 y 9). Para cada condición se tomaron muestras a 3 tiempos distintos: a las 2 horas de incubación con las NPM (carriles 1, 4 y 7), a las 5 horas (carriles 2, 5 y 8) y tras una incubación durante toda la noche (carriles 3, 6 y 9). En todos los ensayos se usó una cantidad fija de lectina (40 µg) que se incubó con las NPM en una relación 1:10. El control de lectina se corresponde con 20 µg totales.

## 11.2. La unión a la agarosa es direccional y mantiene la actividad de las proteínas inmovilizadas

Una vez demostrada que la unión a NPA de los módulos lectina tiene lugar a través de los sitios de unión de azúcares, el siguiente paso consistió en caracterizar la inmovilización de proteínas de fusión. En primer lugar, se empleó la proteína modelo LSLt-EGFP para estu-

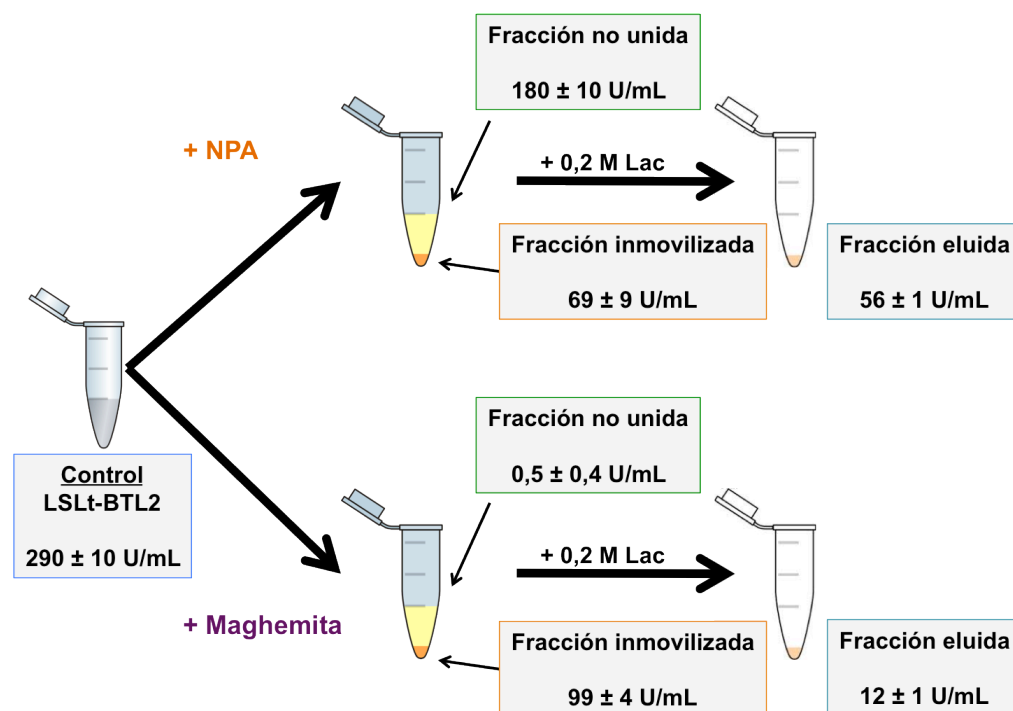
diar la orientación de la proteína de fusión sobre la superficie de las NPA. En este sentido, nos planteamos llevar a cabo una digestión *in situ* con la endoproteasa TEV de la proteína LSLt-EGFP inmovilizada. Si la orientación de la proteína de fusión viene dada por la interacción entre LSLt y la agarosa, el sitio de reconocimiento de la proteasa estará accesible y, por tanto, se podrá llevar a cabo la reacción de digestión, liberándose al medio la proteína uEGFP. Sin embargo, si ésta adoptase algún tipo de conformación que ocultara el sitio de reconocimiento, la digestión no se llevaría a cabo. En el gel de la **Figura 7.33** se muestra el progreso de la digestión enzimática. En primer lugar, se observa cómo a medida que transcurre el tiempo de la reacción, la bandas correspondientes a uEGFP y LSLt crecen en intensidad, mientras que la de LSLt-EGFP se hace menos visible. Esto indica que la reacción progresa adecuadamente y que, por tanto, la proteasa está reconociendo el sitio de corte, lo que supone una correcta orientación de la proteína de fusión sobre la superficie de las NPA. Cuando sólo se estudia la fracción soluble, sin revertir la unión con lactosa, la banda que se observa principalmente en el gel es la correspondiente a uEGFP, lo que indica que el módulo lectina continúa interaccionando con las NPA después de la digestión.



**Figura 7.33.** Ensayo de digestión enzimática con la endoproteasa TEV de la proteína de fusión LSLt-EGFP inmovilizada sobre la superficie de NPA. Carril 1: fracción correspondiente a la proteína LSLt-EGFP unida a NPA. Carril 2: fracción correspondiente a LSLt-EGFP unido a NPA después de 1 h de digestión. Carril 3: fracción correspondiente a LSLt-EGFP unido a NPA después de ~ 20 h de digestión. Carril 4: sobrenadante correspondiente a LSLt-EGFP unido a NPA después de 1 h de digestión. Carril 5: sobrenadante correspondiente a LSLt-EGFP unido a NPA después de ~ 20 h de digestión. Carril 6: incubación con núcleos de maghemita del sobrenadante del carril 5. Carril 7: incubación con NPA del sobrenadante y posterior elución con lactosa del carril 5.

Con la fracción soluble, es decir, con la fracción que no está unida a las NPA tras la digestión con la proteasa TEV, se llevaron a cabo dos ensayos para demostrar que EGFP no interacciona con las NPM de forma inespecífica. Por un lado, esta fracción se incubó durante 2 h a temperatura ambiente con los núcleos de maghemita (*carril 6* del gel de la **Figura**

7.33), obteniéndose un resultado que demuestra que uEGFP no experimenta ningún tipo de unión. De hecho, se observa que desaparece la banda acompañante correspondiente a LSLt, probablemente, debido a la unión inespecífica que ya se había observado previamente con este módulo lectina. En cambio, cuando la fracción soluble se incubaba con NPA (carril 7 del gel de la **Figura 7.33**), resulta que nuevamente solo interacciona LSLt. De modo que queda demostrado que la proteína uEGFP no es reconocida por las nanopartículas, lo que justifica que la proteína de fusión no adopte conformaciones que pudiesen esconder el sitio de reconocimiento de la proteasa TEV.



**Figura 7.34.** Ensayo de actividad enzimática de la proteína de fusión LSLt-BTL2 inmovilizada sobre las NPM recubiertas con agarosa (arriba) y sobre los núcleos de maghemita sin recubrimiento (abajo). La proteína se incubó con las NPM durante 2 h a temperatura ambiente. La fracción inmovilizada se separó de la fracción soluble, donde permanece la proteína que no se ha unido. Posteriormente, la fracción inmovilizada se trató con 0,2 M lactosa para disociar el complejo lectina:NPM y obtener la fracción eluida. En cada una de las fracciones se midió la actividad hidrolítica de la enzima BTL2 según el ensayo descrito en la sección 3.9 de *Materiales y Métodos*.

Usando como modelo la proteína de fusión LSLt-BTL2, se investigó la actividad hidrolítica de la enzima una vez se hubo inmovilizado sobre las NPA o sobre los núcleos de maghemita (**Figura 7.34**). La comparación de las actividades de las fracciones inmovilizadas en sendos soportes muestra que las NPM sin recubrimiento presentan mayor actividad enzimática. Sin embargo, si se compara la actividad de las fracciones no unidas, se observa que apenas hay actividad enzimática en el caso de los núcleos de maghemita, lo que indicaría que prácticamente toda la proteína ha sido retenida por las nanopartículas. Si esto es así, la unión a la maghemita realmente implica una pérdida de actividad específica, en tanto

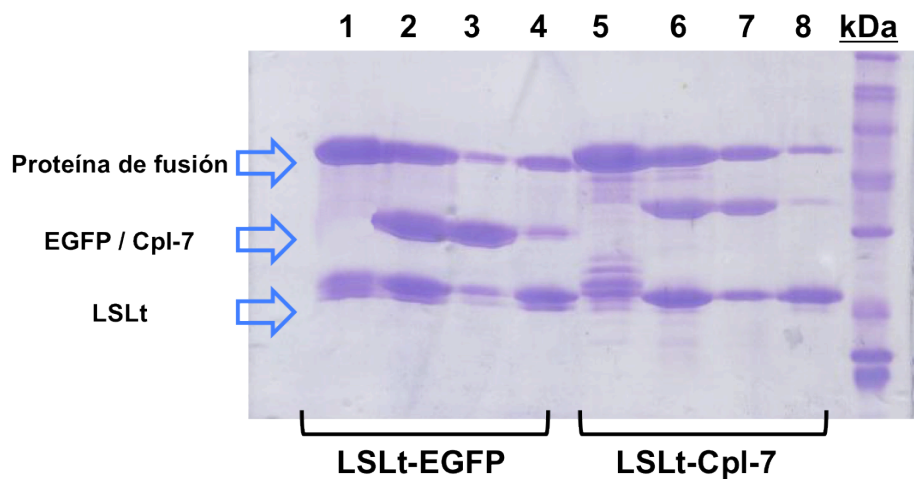
que la actividad registrada en la fracción inmovilizada es mucho menor que la detectada en el control. Por el contrario, aunque la fracción inmovilizada sobre las NPA presenta menor actividad, la suma de las actividades de ambas fracciones (unida y no unida) restituyen la actividad del control. Por lo tanto, aunque la capacidad de carga de las NPA es menor, la enzima inmovilizada mantiene la actividad enzimática, lo que sugiere que la orientación es adecuada y que la accesibilidad del sustrato no es un problema. Asimismo, cuando las fracciones inmovilizadas se tratan con lactosa, solamente la fracción disociada de las NPA posee una actividad similar a la determinada durante la unión. En el caso de los núcleos de maghemita, la actividad de la muestra disociada es mucho menor, lo que posiblemente se deba al hecho de que la lactosa no consigue revertir la unión por completo al tratarse de una unión inespecífica.

Ambos ensayos, tanto el realizado con LSLt-EGFP como el de LSLt-BTL2, demuestran que el módulo lectina permite a las proteínas de fusión adoptar una orientación óptima para que puedan llevar a cabo su función. Esta característica permite englobar a la etiqueta LSLt dentro del grupo de adaptadores basados en biomoléculas, donde se encuentran otros sistemas como el de avidina-biotina o el de reconocimiento de la región Fc de los anticuerpos por parte de proteínas específicas (proteína A y/o G) (Montenegro *et al.*, 2013). Generalmente, este tipo de unión a las NPM es mucho más fuerte que la unión por simple adsorción, aunque cada una de las estrategias tiene sus ventajas y desventajas. Por ejemplo, la principal ventaja de LSLt frente al sistema de avidina-biotina es que no es necesario modificar covalentemente ningún residuo de la proteína, lo que podría afectar negativamente a la actividad de ciertas proteínas. Además, al igual que la unión de la avidina con la biotina, la interacción de la agarosa con LSLt es muy fuerte gracias a la multivalencia y resiste condiciones químicas duras, incluyendo un amplio rango de pH y el uso de solventes orgánicos (López-Gallego *et al.*, 2012). Detalles tales como la orientación o la fuerza de la unión con la NPM son muy importantes para el desarrollo de futuras aplicaciones (por ejemplo, biosensores) basadas en el uso de LSLt como sistema de anclaje a NPA.

### **11.3. Las nanopartículas como un sistema de limpieza**

Con la finalidad de optimizar el sistema de expresión-purificación de proteínas recombinantes empleando el vector pKLSLt, hemos desarrollado un sistema complementario basado en el uso de las NPA. Las nanopartículas serían utilizadas como un sistema de limpieza para poder eliminar la etiqueta de fusión generada de forma espontánea y retirar simultáneamente la endoproteasa del medio una vez transcurrida la digestión, evitando un paso

cromatográfico durante la purificación. El primer paso para implementar este sistema consistió en generar una proteasa TEV fusionada a LSLt (el vector pKLSLt se modificó para tal fin, cambiando la secuencia de reconocimiento de la endoproteasa TEV para evitar el auto-corte). La proteína de fusión LSLt-TEV se produjo con un buen rendimiento (~6 mg/L de cultivo) y se purificó normalmente. Asimismo, se comprobó que era activa en tampón TSA en una relación 1:80 (TEV:proteína de fusión) frente a las proteínas LSLt-EGFP y LSLt-Cpl-7 (**Figura 7.35**), al igual que la versión con la cola de histidinas, utilizada hasta este momento. Cuando los diferidos se incubaron con las NPA durante 2 h a temperatura ambiente, la proteasa, la etiqueta LSLt y la proteína de fusión que quedaba sin digerir se unieron a las nanopartículas, quedando en solución principalmente la proteína diana (**Figura 7.35**). A la vista del gel, la cantidad de nanopartículas debe ajustarse para poder retirar por completo todas las trazas de etiqueta de fusión y proteína de fusión que quedan en el medio y conseguir un alto nivel de pureza. Tras el lavado con las NPA, la proteína puede someterse al paso por una cromatografía de exclusión por tamaño para eliminar los agregados solubles y conseguir mayor homogeneidad conformacional. La ventaja de este sistema es que las partículas son reutilizables, puesto que con un lavado con una solución con lactosa a alta concentración (0,1–0,4 M), las NPA se descargan y vuelven a estar listas para ser utilizadas de nuevo.



**Figura 7.35.** Digestión de las proteínas de fusión LSLt-EGFP y LSLt-Cpl-7 con la proteasa LSLt-TEV e incubación con las NPA. Carriles 1 y 5: proteína de fusión sin digerir. Carriles 2 y 6: digerido de las proteínas de fusión con LSLt-TEV. Carriles 3 y 7: fracción no unida después de incubar con NPA durante 2 h a temperatura ambiente. Carriles 4 y 8: fracción eluida al tratar las NPA con 0,1 M lactosa durante 1 h a temperatura ambiente.

# CAPÍTULO 8



## CONCLUSIONES



## VIII

## CONCLUSIONES

## 1. Obtención y caracterización de lectinas fúngicas

- 1.1. El módulo lectina N-terminal de la proteína LSLa del hongo basidiomiceto *Laetiporus sulphureus* (LSL<sub>150</sub>) es una unidad de plegamiento independiente. Se ha comprobado que actúa como una etiqueta de solubilización natural para el respectivo módulo C-terminal formador de poros (PFM).
- 1.2. La estructura cristalográfica de LSL<sub>150</sub> revela un plegamiento de tipo trébol  $\beta$  equivalente al descrito previamente para el módulo lectina en la proteína completa. El estudio detallado del complejo LSL<sub>150</sub>:(lactosa) $\beta\gamma$  ha permitido identificar dos sitios de unión diferentes en los subdominios  $\beta$  y  $\gamma$  y caracterizar así las interacciones que se establecen entre la proteína, el ligando y las moléculas de solvente. El análisis termodinámico de la unión de lactosa permitió clasificar los dos sitios de unión en uno de alta afinidad relativa (sitio  $\gamma$ ) y otro de baja (sitio  $\beta$ ).
- 1.3. La comparación de las estructuras de LSL<sub>150</sub> con lactosas y sin ellas ha permitido dibujar un mapa del reordenamiento del solvente durante el reconocimiento del ligando.
- 1.4. Se ha clonado el gen *lbl* del hongo basidiomiceto *Laccaria bicolor* y se ha demostrado que posee una secuencia diferente a la predicha. Además, se ha observado que el gen *lbl* posee mayores niveles de expresión en los estadios tempranos de crecimiento de los cuerpos fructíferos del hongo.
- 1.5. La proteína completa LBL no se produce de forma soluble recombinantemente ni en *E. coli* ni en *S. cerevisiae*. Sin embargo, el módulo lectina N-terminal de la proteína (LBL<sub>152</sub>) se expresa y purifica con un rendimiento muy elevado, comparable al de LSL<sub>150</sub>.



- 1.6. La estructura tridimensional de LBL<sub>152</sub> en complejo con lactosa ha confirmado que la proteína LBL es un nuevo miembro de la familia de los tréboles  $\beta$ . Esta proteína tiene capacidad para unir tres moléculas de ligando, una por subdominio. La resolución atómica del modelo ( $\sim 1$  Å) ha permitido caracterizar en detalle las interacciones, donde las moléculas de agua juegan un papel fundamental.
- 1.7. La caracterización termodinámica de la unión de lactosa indica que dos de los tres sitios de unión son de alta afinidad relativa. Estos resultados junto con la información obtenida del modelo cristalográfico indican que el sitio  $\alpha$  sería uno de los sitios de mayor afinidad.
- 1.8. La comparación de las estructuras cristalográficas de LSL<sub>150</sub> y LBL<sub>152</sub> confirma que las moléculas de agua forman parte de la arquitectura de los sitios de unión.
- 1.9. Estudios mediante espectroscopía de fluorescencia en estado estacionario han demostrado que la estructura de LBL<sub>152</sub> sin lactosa es esencialmente la misma que la del complejo LBL<sub>152</sub>:(lactosa) $\alpha\beta\gamma$ , aunque existen cambios locales en el entorno de los sitios de unión producidos por la unión del ligando.
- 1.10. Las curvas de desnaturalización térmica de LSL<sub>150</sub> y LBL<sub>152</sub> muestran un incremento considerable de la estabilidad térmica promovido por la unión de lactosa.
- 1.11. Las lectinas SSL de *Sclerotinia sclerotiorum* y ANL de *Aspergillus niger* se producen con muy bajo rendimiento de forma recombinante. Los modelos obtenidos por homología indican que SSL sería un nuevo trébol  $\beta$  con un único sitio de unión, en tanto que ANL, también un trébol  $\beta$ , presentaría dos sitios de unión independientes.

## 2. Estudio de variantes de LSL<sub>150</sub>

- 2.1. Mutaciones puntuales para reconstruir un sitio  $\gamma$  en el subdominio  $\alpha$  (S29H, S37H, A43F) no consiguen reconstituir la capacidad para unir azúcares. Sin embargo, los mutantes triples que incorporan tres mutaciones (LSL<sub>150</sub>- $\alpha$ F y LSL<sub>150</sub>- $\alpha$ Y) sí que lo consiguen. El estudio detallado de las estructuras cristalográficas de todos los mutantes advierte del papel clave de las moléculas de agua como determinantes de la unión.

- 2.2. A pesar de que la estructura global de la lectina no se ve modificada en los triples mutantes, la adquisición de un nuevo sitio de unión (más funcionalidad) va asociada a una pérdida de la estabilidad térmica, tal y como se desprende de las curvas de desnaturalización térmica obtenidas mediante dicroísmo circular.
- 2.3. El nuevo sitio de unión  $\alpha$  presenta una afinidad intermedia entre los sitios  $\beta$  y  $\gamma$  en ambos mutantes triples.
- 2.4. Se ha identificado un nuevo ligando (3'-galactosillactosa) en el sitio  $\gamma$  del mutante triple LSL<sub>150</sub>- $\alpha$ F. El nuevo ligando posee un residuo de galactosa unido al anillo de galactosa de la lactosa mediante enlace  $\beta(1\rightarrow3)$  y establece nuevas interacciones con los residuos del sitio de unión.
- 2.5. Los intentos por conseguir un sitio  $\beta$  de mayor afinidad en LSL<sub>150</sub> han sido insatisfactorios. Uno de los mutantes (P77H) no pudo ser caracterizado y el otro (I85H) muestra como el residuo mutado adopta una conformación incompatible con el modo de reconocimiento de los azúcares.

### 3. Análisis funcional de la unión a azúcares

- 3.1. Los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> presentan un perfil similar de unión de azúcares. Ambos módulos tienen predilección por oligosacáridos ricos en poli-LacNAc, estructuras con cierta relevancia biológica, sobre todo para los glicoepítomos del sistema sanguíneo.
- 3.2. El análisis estructural de los complejos de LSL<sub>150</sub> y LBL<sub>152</sub> en presencia de LacNAc indica que ambas lectinas establecen nuevas interacciones que involucran al grupo acetamido del anillo de glucosa, lo que confiere especificidad. Además, los ensayos calorimétricos corroboran que LacNAc se une y que lo hace con mayor afinidad que la observada para la lactosa.
- 3.3. Estudios de competición utilizando glicoproteínas inmovilizadas y una colección de azúcares en solución muestran que los módulos LSL<sub>150</sub> y LBL<sub>152</sub> (y, por lo tanto, las proteínas LSLa y LBL) pueden reconocer glicoepítomos de insectos y nematodos. Además,

éstas reconocen extractos celulares de *C. elegans* y *D. melanogaster* de forma específica, lo cual podría dar pistas acerca de su función.

## 4. Aplicaciones biotecnológicas novedosas

- 4.1. Los módulos lectina LSL<sub>150</sub> y LBL<sub>152</sub> actúan como etiquetas de fusión para la producción de proteínas recombinantes. En este sentido, se ha diseñado un set de plásmidos que permiten la clonación y expresión de genes y la posterior purificación en un único paso de proteínas recombinantes mediante un sistema de cromatografía de afinidad a columnas de agarosa. Las proteínas de fusión se producen etiquetadas en el extremo N-terminal y separadas de la etiqueta con un espaciador (ASSS) y un sitio de reconocimiento de la endoproteasa TEV (ENLYFQG) que permite la liberación del módulo lectina.
- 4.2. El módulo lectina en las proteínas de fusión queda orientado de tal manera que no influye *a priori* ni en el plegamiento ni en la actividad de la proteína acompañante, como se ha demostrado en este trabajo con las proteínas de fusión LSLt-EGFP, LSLt-Bgl y LSLt-BTL2.
- 4.3. Se ha desarrollado un nuevo método de inmovilización reversible de enzimas mediante la fijación a columnas de agarosa a través del módulo lectina LSL<sub>150</sub>. La proteína inmovilizada se distribuye homogéneamente en la resina, lo que potencia su uso en diferentes aplicaciones de interés industrial.
- 4.4. La inmovilización a través del módulo lectina es estable en el tiempo y aguanta diferentes condiciones extremas, como el uso de disolventes orgánicos, pH ácidos y elevadas temperaturas.
- 4.5. Se ha desarrollado un nuevo método de inmovilización de proteínas a través del módulo lectina LSL<sub>150</sub> a nanopartículas magnéticas recubiertas con agarosa.
- 4.6. La unión a las nanopartículas magnéticas es reversible, direccional y mantiene la actividad de las proteínas una vez se han inmovilizado.
- 4.7. Las nanopartículas magnéticas se pueden utilizar como un sistema de limpieza de reacciones en las que intervienen proteínas etiquetadas con los módulos lectina LSL<sub>150</sub> y/o LBL<sub>152</sub>.

# BIBLIOGRAFÍA





Abby, S. y Daubin, V. (2007). *Comparative Genomics and the Evolution of Prokaryotes*. Trends Microbiol **15**: 135-141.

Acebrón, I., Curiel, J.A., de las Rivas, B., Muñoz, R. y Mancheño, J.M. (2009). *Cloning, Production, Purification and Preliminary Crystallographic Analysis of a Glycosidase from the Food Lactic Acid Bacterium Lactobacillus Plantarum Cect 748(T)*. Protein Expr Purif **68**: 177-182.

Adair, W.L. y Kornfeld, S. (1974). *Isolation of the Receptors for Wheat Germ Agglutinin and the Ricinus Communis Lectins from Human Erythrocytes Using Affinity Chromatography*. J Biol Chem **249**: 4696-4704.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. y Zwart, P.H. (2010). *Phenix: A Comprehensive Python-Based System for Macromolecular Structure Solution*. Acta Crystallogr D Biol Crystallogr **66**: 213-221.

Aebi, M. (2013). *N-Linked Protein Glycosylation in the Er*. Biochim Biophys Acta **1833**: 2430-2437.

Ahmed, H., Du, S.J., O'leary, N. y Vasta, G.R. (2004). *Biochemical and Molecular Characterization of Galectins from Zebrafish (Danio Rerio): Notochord-Specific Expression of a Prototype Galectin During Early Embryogenesis*. Glycobiology **14**: 219-232.

Álvarez, C.P., Lasala, F., Carrillo, J., Muñiz, O., Corbi, A.L. y Delgado, R. (2002). *C-Type Lectins Dc-Sign and L-Sign Mediate Cellular Entry by Ebola Virus in Cis and in Trans*. J Virol **76**: 6841-6844.

Amselem, J., Cuomo, C.A., Van Kan, J.A., Viaud, M., Benito, E.P., Couloux, A., Coutinho, P.M., De Vries, R.P., Dyer, P.S., Fillinger, S., Fournier, E., Gout, L., Hahn, M., Kohn, L., Lapalu, N., Plummer, K.M., Pradier, J.M., Quevillon, E., Sharon, A., Simon, A., Ten Have, A., Tudzynski, B., Tudzynski, P., Wincker, P., Andrew, M., Anthouard, V., Beever, R.E., Beffa, R., Benoit, I., Bouzid, O., Brault, B., Chen, Z., Choquer, M., Collemare, J., Cotton, P., Danchin, E.G., Da Silva, C., Gautier, A., Giraud, C., Giraud, T., Gonzalez, C., Grossetete, S., Guldener, U., Henrissat, B., Howlett, B.J., Kodira, C., Kretschmer, M., Lappartient, A., Leroch, M., Levis, C., Mauceli, E., Neuveglise, C., Oeser, B., Pearson, M., Poulain, J., Poussereau, N., Quesneville, H., Rasclé, C., Schumacher, J., Segurens, B., Sexton, A., Silva, E., Sirven, C., Soanes, D.M., Talbot, N.J., Templeton, M., Yandava, C., Yarden, O., Zeng, Q., Rollins, J.A., Lebrun, M.H. y Dickman, M. (2011). *Genomic Analysis of the Necrotrophic Fungal Pathogens Sclerotinia Sclerotiorum and Botrytis Cinerea*. PLoS Genet **7**: e1002230.

Angulo, I., Acebrón, I., De Las Rivas, B., Munoz, R., Rodriguez-Crespo, I., Menendez, M., Garcia, P., Tateno, H., Goldstein, I.J., Perez-Agote, B. y Mancheno, J.M. (2011). *High-Resolution Structural Insights on the Sugar-Recognition and Fusion Tag Properties of a Versatile Beta-Trefoil Lectin Domain from the Mushroom Laetiporus Sulphureus*. Glycobiology **21**: 1349-1361.

Arnau, J., Lauritzen, C., Petersen, G.E. y Pedersen, J. (2006). *Current Strategies for the Use of Affinity Tags and Tag Removal for the Purification of Recombinant Proteins*. Protein Expr Purif **48**: 1-13.

Arnold, K., Bordoli, L., Kopp, J. y Schwede, T. (2006). *The Swiss-Model Workspace: A Web-Based Environment for Protein Structure Homology Modelling*. Bioinformatics **22**: 195-201.

- Babál, P., Pindak, F.F., Wells, D.J. y Gardner, W.A., Jr. (1994). *Purification and Characterization of a Sialic Acid-Specific Lectin from Tritrichomonas Mobilensis*. *Biochem J* **299** ( Pt 2): 341-346.
- Baños-Sanz, J.I., Villate, M., Sanz-Aparicio, J., Brearley, C.A. y González, B. (2010). *Crystallization and Preliminary X-Ray Diffraction Analysis of Inositol 1,3,4,5,6-Pentakisphosphate Kinase from Arabidopsis Thaliana*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**: 102-106.
- Bertozzi, C.R. y Kiessling, L.L. (2001). *Chemical Glycobiology*. *Science* **291**: 2357-2364.
- Bhatia, Y., Mishra, S. y Bisaria, V.S. (2002). *Microbial Beta-Glucosidases: Cloning, Properties, and Applications*. *Crit Rev Biotechnol* **22**: 375-407.
- Bhowal, J., Guha, A.K. y Chatterjee, B.P. (2005). *Purification and Molecular Characterization of a Sialic Acid Specific Lectin from the Phytopathogenic Fungus Macrophomina Phaseolina*. *Carbohydr Res* **340**: 1973-1982.
- Birck, C., Damian, L., Marty-Detraves, C., Lougarre, A., Schulze-Briese, C., Koehl, P., Fournier, D., Paquereau, L. y Samama, J.P. (2004). *A New Lectin Family with Structure Similarity to Actinoporins Revealed by the Crystal Structure of Xerocomus Chrysenteron Lectin Xcl*. *J Mol Biol* **344**: 1409-1420.
- Bleuler-Martínez, S., Butschi, A., Garbani, M., Walti, M.A., Wohlschlager, T., Potthoff, E., Sabotic, J., Pohleven, J., Luthy, P., Hengartner, M.O., Aebi, M. y Kunzler, M. (2011). *A Lectin-Mediated Resistance of Higher Fungi against Predators and Parasites*. *Mol Ecol* **20**: 3056-3070.
- Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M.E., Alvarez, R., Bryan, M.C., Fazio, F., Calarese, D., Stevens, J., Razi, N., Stevens, D.J., Skehel, J.J., Van Die, I., Burton, D.R., Wilson, I.A., Cummings, R., Bovin, N., Wong, C.H. y Paulson, J.C. (2004). *Printed Covalent Glycan Array for Ligand Profiling of Diverse Glycan Binding Proteins*. *Proc Natl Acad Sci U S A* **101**: 17033-17038.
- Bojarova, P. y Kren, V. (2009). *Glycosidases: A Key to Tailored Carbohydrates*. *Trends Biotechnol* **27**: 199-209.
- Bolivar, J.M., Hidalgo, A., Sánchez-Ruiloba, L., Berenguer, J., Guisán, J.M. y López-Gallego, F. (2011). *Modulation of the Distribution of Small Proteins within Porous Matrixes by Smart-Control of the Immobilization Rate*. *J Biotechnol* **155**: 412-420.
- Boraston, A.B., Bolam, D.N., Gilbert, H.J. y Davies, G.J. (2004). *Carbohydrate-Binding Modules: Fine-Tuning Polysaccharide Recognition*. *Biochem J* **382**: 769-781.
- Borchert, A. y Buchholz, K. (1984). *Improved Biocatalyst Effectiveness by Controlled Immobilization of Enzymes*. *Biotechnol Bioeng* **26**: 727-736.
- Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. y Schwede, T. (2009). *Protein Structure Homology Modeling Using Swiss-Model Workspace*. *Nat Protoc* **4**: 1-13.
- Bourenkov, G.P. y Popov, A.N. (2006). *A Quantitative Approach to Data-Collection Strategies*. *Acta Crystallogr D Biol Crystallogr* **62**: 58-64.

- Bovi, M., Cenci, L., Perduca, M., Capaldi, S., Carrizo, M.E., Civiero, L., Chiarelli, L.R., Galliano, M. y Monaco, H.L. (2013). *Bel Beta-Trefoil: A Novel Lectin with Antineoplastic Properties in King Bolete (Boletus Edulis) Mushrooms*. *Glycobiology* **23**: 578-592.
- Bradford, M.M. (1976). *A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding*. *Anal Biochem* **72**: 248-254.
- Brahms, S. y Brahms, J. (1980). *Determination of Protein Secondary Structure in Solution by Vacuum Ultraviolet Circular Dichroism*. *J Mol Biol* **138**: 149-178.
- Braud, S., Moutiez, M., Belin, P., Abello, N., Drevet, P., Zinn-Justin, S., Courcon, M., Masson, C., Dassa, J., Charbonnier, J.B., Boulain, J.C., Menez, A., Genet, R. y Gondry, M. (2005). *Dual Expression System Suitable for High-Throughput Fluorescence-Based Screening and Production of Soluble Proteins*. *J Proteome Res* **4**: 2137-2147.
- Brent, M.R. (2008). *Steady Progress and Recent Breakthroughs in the Accuracy of Automated Genome Annotation*. *Nat Rev Genet* **9**: 62-73.
- Brewer, C.F., Miceli, M.C. y Baum, L.G. (2002). *Clusters, Bundles, Arrays and Lattices: Novel Mechanisms for Lectin-Saccharide-Mediated Cellular Interactions*. *Curr Opin Struct Biol* **12**: 616-623.
- Broom, A., Doxey, A.C., Lobsanov, Y.D., Berthin, L.G., Rose, D.R., Howell, P.L., Mcconkey, B.J. y Meiering, E.M. (2012). *Modular Evolution and the Origins of Symmetry: Reconstruction of a Three-Fold Symmetric Globular Protein*. *Structure* **20**: 161-171.
- Brych, S.R., Blaber, S.I., Logan, T.M. y Blaber, M. (2001). *Structure and Stability Effects of Mutations Designed to Increase the Primary Sequence Symmetry within the Core Region of a Beta-Trefoil*. *Protein Sci* **10**: 2587-2599.
- Burnette, W.N. (1981). *"Western Blotting": Electrophoretic Transfer of Proteins from Sodium Dodecyl Sulfate--Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A*. *Anal Biochem* **112**: 195-203.
- Camby, I., Le Mercier, M., Lefranc, F. y Kiss, R. (2006). *Galectin-1: A Small Protein with Major Functions*. *Glycobiology* **16**: 137R-157R.
- Campanero-Rhodes, M.A., Menéndez, M., Saiz, J.L., Sanz, L., Calvete, J.J. y Solís, D. (2005). *Analysis of the Stability of the Spermadhesin Psp-I/Psp-Ii Heterodimer. Effects of Zn<sup>2+</sup> and Acidic Ph*. *FEBS J* **272**: 5663-5670.
- Candy, L., Van Damme, E.J., Peumans, W.J., Menu-Bouaouiche, L., Erard, M. y Rouge, P. (2003). *Structural and Functional Characterization of the Galnac/Gal-Specific Lectin from the Phytopathogenic Ascomycete Sclerotinia Sclerotiorum (Lib.) De Bary*. *Biochem Biophys Res Commun* **308**: 396-402.
- Carlie, M.J., Watkinson, S.C. y Gooday, G.W. (2001). *The Fungi*, Academic Press, Londres



- Carrasco-López, C., Godoy, C., de las Rivas, B., Fernández-Lorente, G., Palomo, J.M., Guisán, J.M., Fernández-Lafuente, R., Martínez-Ripoll, M. y Hermoso, J.A. (2008). *Crystallization and Preliminary X-Ray Diffraction Studies of the Btl2 Lipase from the Extremophilic Microorganism Bacillus Thermocatenulatus*. Acta Crystallogr Sect F Struct Biol Cryst Commun **64**: 1043-1045.
- Carrington, J.C. y Dougherty, W.G. (1987). *Small Nuclear Inclusion Protein Encoded by a Plant Potyvirus Genome Is a Protease*. J Virol **61**: 2540-2548.
- Carrizo, M.E., Capaldi, S., Perduca, M., Irazoqui, F.J., Nores, G.A. y Monaco, H.L. (2005). *The Antineoplastic Lectin of the Common Edible Mushroom (Agaricus Bisporus) Has Two Binding Sites, Each Specific for a Different Configuration at a Single Epimeric Hydroxyl*. J Biol Chem **280**: 10614-10623.
- Cavada, B.S., Santos, C.F., Grangeiro, T.B., Nunes, E.P., Sales, P.V., Ramos, R.L., De Sousa, F.A., Crisostomo, C.V. y Calvete, J.J. (1998). *Purification and Characterization of a Lectin from Seeds of Vatairea Macrocarpa Duke*. Phytochemistry **49**: 675-680.
- Cavalier-Smith, T. (1998). *A Revised Six-Kingdom System of Life*. Biol Rev Camb Philos Soc **73**: 203-266.
- Cavalier-Smith, T. (2004). *Only Six Kingdoms of Life*. Proc Biol Sci **271**: 1251-1262.
- Cerigini, E., Palma, F., Barbieri, E., Buffalini, M. y Stocchi, V. (2008). *The Tuber Borchii Fruiting Body-Specific Protein Tbf-1, a Novel Lectin Which Interacts with Associated Rhizobium Species*. FEMS Microbiol Lett **284**: 197-203.
- Charcot, J.M. y Robin, C. (1853). *Observation De Leucocythemie*. C R Mem Soc Biol **5**: 44-50.
- Chumkhunthod, P., Rodtong, S., Lambert, S.J., Fordham-Skelton, A.P., Rizkallah, P.J., Wilkinson, M.C. y Reynolds, C.D. (2006). *Purification and Characterization of an N-Acetyl-D-Galactosamine-Specific Lectin from the Edible Mushroom Schizophyllum Commune*. Biochim Biophys Acta **1760**: 326-332.
- Clancy, S. (2008). *Rna Splicing: Introns, Exons and Spliceosome*, Nature Education.
- Corpet, F. (1988). *Multiple Sequence Alignment with Hierarchical Clustering*. Nucleic Acids Res **16**: 10881-10890.
- Cowtan, K., Emsley, P. y Wilson, K.S. (2011). *From Crystal to Structure with Ccp4*. Acta Crystallogr D Biol Crystallogr **67**: 233-234.
- Cummings, R.D. y Pierce, J.M. (2014). *The Challenge and Promise of Glycomics*. Chem Biol **21**: 1-15.
- Curiel, J.A., De Las Rivas, B., Mancheño, J.M. y Muñoz, R. (2011). *The Puri Family of Expression Vectors: A Versatile Set of Ligation Independent Cloning Plasmids for Producing Recombinant His-Fusion Proteins*. Protein Expr Purif **76**: 44-53.

- Dahms, N.M., Lobel, P. y Kornfeld, S. (1989). *Mannose 6-Phosphate Receptors and Lysosomal Enzyme Targeting*. J Biol Chem **264**: 12115-12118.
- Dai, Z., Kawde, A.N., Xiang, Y., La Belle, J.T., Gerlach, J., Bhavanandan, V.P., Joshi, L. y Wang, J. (2006). *Nanoparticle-Based Sensing of Glycan-Lectin Interactions*. J Am Chem Soc **128**: 10018-10019.
- Dam, T.K. y Brewer, C.F. (2002). *Thermodynamic Studies of Lectin-Carbohydrate Interactions by Isothermal Titration Calorimetry*. Chem Rev **102**: 387-429.
- Dam, T.K. y Brewer, C.F. (2010). *Lectins as Pattern Recognition Molecules: The Effects of Epitope Density in Innate Immunity*. Glycobiology **20**: 270-279.
- Dam, T.K., Roy, R., Page, D. y Brewer, C.F. (2002). *Negative Cooperativity Associated with Binding of Multivalent Carbohydrates to Lectins. Thermodynamic Analysis of the "Multivalency Effect"*. Biochemistry **41**: 1351-1358.
- Davis, G.D., Elisee, C., Newham, D.M. y Harrison, R.G. (1999). *New Fusion Protein Systems Designed to Give Soluble Expression in Escherichia Coli*. Biotechnol Bioeng **65**: 382-388.
- Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S. y Richardson, D.C. (2007). *Molprobit: All-Atom Contacts and Structure Validation for Proteins and Nucleic Acids*. Nucleic Acids Res **35**: W375-383.
- de las Rivas, B., Curiel, J.A., Mancheño, J.M. y Muñoz, R. (2007). *Expression Vectors for Enzyme Restriction- and Ligation-Independent Cloning for Producing Recombinant His-Fusion Proteins*. Biotechnol Prog **23**: 680-686.
- Derewenda, Z., Yariv, J., Helliwell, J.R., Kalb, A.J., Dodson, E.J., Papiz, M.Z., Wan, T. y Campbell, J. (1989). *The Structure of the Saccharide-Binding Site of Concanavalin A*. EMBO J **8**: 2189-2193.
- Di Guan, C., Li, P., Riggs, P.D. y Inouye, H. (1988). *Vectors That Facilitate the Expression and Purification of Foreign Peptides in Escherichia Coli by Fusion to Maltose-Binding Protein*. Gene **67**: 21-30.
- Di Lella, S., Marti, M.A., Alvarez, R.M., Estrin, D.A. y Ricci, J.C. (2007). *Characterization of the Galectin-1 Carbohydrate Recognition Domain in Terms of Solvent Occupancy*. J Phys Chem B **111**: 7360-7366.
- Digabriele, A.D., Lax, I., Chen, D.I., Svahn, C.M., Jaye, M., Schlessinger, J. y Hendrickson, W.A. (1998). *Structure of a Heparin-Linked Biologically Active Dimer of Fibroblast Growth Factor*. Nature **393**: 812-817.
- Dobson, J. (2006). *Magnetic Nanoparticles for Drug Delivery*. Drug Development Research **67**: 55-60.
- Dodd, R.B. y Drickamer, K. (2001). *Lectin-Like Proteins in Model Organisms: Implications for Evolution of Carbohydrate-Binding Activity*. Glycobiology **11**: 71R-79R.
- Drickamer, K. (1988). *Two Distinct Classes of Carbohydrate-Recognition Domains in Animal Lectins*. J Biol Chem **263**: 9557-9560.

Drickamer, K. (1997). *Making a Fitting Choice: Common Aspects of Sugar-Binding Sites in Plant and Animal Lectins*. *Structure* **5**: 465-468.

Drickamer, K. (1999). *C-Type Lectin-Like Domains*. *Curr Opin Struct Biol* **9**: 585-590.

Dummler, A., Lawrence, A.M. y De Marco, A. (2005). *Simplified Screening for the Detection of Soluble Fusion Constructs Expressed in E. Coli Using a Modular Set of Vectors*. *Microb Cell Fact* **4**: 34.

Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W. y Hofnung, M. (1988). *Sequence of Male Gene and of Its Product, the Maltose-Binding Protein of Escherchia Coli K12*. *J. Biol. Chem.* **259**: 10606-10613.

Dyson, M.R., Shadbolt, S.P., Vincent, K.J., Perera, R.L. y Mccafferty, J. (2004). *Production of Soluble Mammalian Proteins in Escherichia Coli: Identification of Protein Features That Correlate with Successful Expression*. *BMC Biotechnol* **4**: 32.

Earl, L.A., Bi, S. y Baum, L.G. (2011). *Galectin Multimerization and Lattice Formation Are Regulated by Linker Region Structure*. *Glycobiology* **21**: 6-12.

Einhauer, A. y Jungbauer, A. (2001). *The Flag Peptide, a Versatile Fusion Tag for the Purification of Recombinant Proteins*. *J Biochem Biophys Methods* **49**: 455-465.

Emsley, P. y Cowtan, K. (2004). *Coot: Model-Building Tools for Molecular Graphics*. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132.

Erjavec, J., Kos, J., Ravnikar, M., Dreö, T. y Sabotic, J. (2012). *Proteins of Higher Fungi--from Forest to Application*. *Trends Biotechnol* **30**: 259-273.

Ermakova, E., Miller, M.C., Nesmelova, I.V., Lopez-Merino, L., Berbis, M.A., Nesmelov, Y., Tkachev, Y.V., Lagartera, L., Daragan, V.A., Andre, S., Canada, F.J., Jimenez-Barbero, J., Solis, D., Gabius, H.J. y Mayo, K.H. (2013). *Lactose Binding to Human Galectin-7 (P53-Induced Gene 1) Induces Long-Range Effects through the Protein Resulting in Increased Dimer Stability and Evidence for Positive Cooperativity*. *Glycobiology* **23**: 508-523.

Esposito, D. y Chatterjee, D.K. (2006). *Enhancement of Soluble Protein Expression through the Use of Fusion Tags*. *Curr Opin Biotechnol* **17**: 353-358.

Evans, P.R. (1999). *Some Notes on Choices in Data Collection*. *Acta Crystallogr D Biol Crystallogr* **55**: 1771-1772.

Evans, P.R. (2006). *Scaling and Assessment of Data Quality*. *Acta Crystallogr D Biol Crystallogr* **62**: 72-82.

Evans, P.R. (2011). *An Introduction to Data Reduction: Space-Group Determination, Scaling and Intensity Statistics*. *Acta Crystallogr D Biol Crystallogr* **67**: 282-292.

Fazio, F., Bryan, M.C., Blixt, O., Paulson, J.C. y Wong, C.H. (2002). *Synthesis of Sugar Arrays in Microtiter Plate*. *J Am Chem Soc* **124**: 14397-14402.

Feng, J., Li, M., Huang, Y. y Xiao, Y. (2010). *Symmetric Key Structural Residues in Symmetric Proteins with Beta-Trefoil Fold*. PLoS One **5**: e14138.

Fernández-Lorente, G., Godoy, C.A., Mendes, A.A., López-Gallego, F., Grazu, V., De Las Rivas, B., Palomo, J.M., Hermoso, J., Fernández-Lafuente, R. y Guisán, J.M. (2008). *Solid-Phase Chemical Amination of a Lipase from Bacillus Thermocatenulatus to Improve Its Stabilization Via Covalent Immobilization on Highly Activated Glyoxyl-Agarose*. Biomacromolecules **9**: 2553-2561.

Fersht, A.R. (1987). *The Hydrogen Bond in Molecular Recognition*. Trends Biochem Sci **12**: 301-304.

Fox, E.M. y Howlett, B.J. (2008). *Secondary Metabolism: Regulation and Role in Fungal Biology*. Curr Opin Microbiol **11**: 481-487.

Fukui, S., Feizi, T., Galustian, C., Lawson, A.M. y Chai, W. (2002). *Oligosaccharide Microarrays for High-Throughput Detection and Specificity Assignments of Carbohydrate-Protein Interactions*. Nat Biotechnol **20**: 1011-1017.

Gaberc-Porekar, V. y Menart, V. (2001). *Perspectives of Immobilized-Metal Affinity Chromatography*. J Biochem Biophys Methods **49**: 335-360.

Gabius, H.J. (2009). *The Sugar Code: Fundamentals of Glycosciences*, Wiley-Blackwell, Weinheim

Gabius, H.J., Andre, S., Jimenez-Barbero, J., Romero, A. y Solís, D. (2011). *From Lectin Structure to Functional Glycomics: Principles of the Sugar Code*. Trends Biochem Sci **36**: 298-313.

Gambaryan, A.S., Tuzikov, A.B., Piskarev, V.E., Yamnikova, S.S., Lvov, D.K., Robertson, J.S., Bovin, N.V. y Matrosovich, M.N. (1997). *Specification of Receptor-Binding Phenotypes of Influenza Virus Isolates from Different Hosts Using Synthetic Sialylglycopolymers: Non-Egg-Adapted Human H1 and H3 Influenza A and Influenza B Viruses Share a Common High Binding Affinity for 6'-Sialyl(N-Acetyl)lactosamine*. Virology **232**: 345-350.

Ganiko, L., Puccia, R., Mariano, V.S., Sant'anna, O.A., Freymuller, E., Roque-Barreira, M.C. y Travassos, L.R. (2007). *Paracoccin, an N-Acetyl-Glucosamine-Binding Lectin of Paracoccidioides Brasiliensis, Is Involved in Fungal Growth*. Microbes Infect **9**: 695-703.

Garcia, P., Garcia, J.L., Garcia, E., Sanchez-Puelles, J.M. y Lopez, R. (1990). *Modular Organization of the Lytic Enzymes of Streptococcus Pneumoniae and Its Bacteriophages*. Gene **86**: 81-88.

García-Segura, J.M., Gavilanes, J.G., Martínez Del Pozo, Á., Montero, F., Oñaderra, M. y Vivanco, F. (2002). *Técnicas Instrumentales De Análisis En Bioquímica*, Editorial Síntesis, Madrid

Garman, E. (1999). *Cool Data: Quantity and Quality*. Acta Crystallogr D Biol Crystallogr **55**: 1641-1653.

Garman, E.F. y Owen, R.L. (2006). *Cryocooling and Radiation Damage in Macromolecular Crystallography*. Acta Crystallogr D Biol Crystallogr **62**: 32-47.

- Garman, S.C. y Garboczi, D.N. (2004). *The Molecular Defect Leading to Fabry Disease: Structure of Human Alpha-Galactosidase*. J Mol Biol **337**: 319-335.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. y Bairoch, A. (2003). *Expasy: The Proteomics Server for in-Depth Protein Knowledge and Analysis*. Nucleic Acids Res **31**: 3784-3788.
- Geiser, M., Cebe, R., Drewello, D. y Schmitz, R. (2001). *Integration of Pcr Fragments at Any Specific Site within Cloning Vectors without the Use of Restriction Enzymes and DNA Ligase*. Biotechniques **31**: 88-90, 92.
- Giblett, E.R. y Crookston, M.C. (1964). *Agglutinability of Red Cells by Anti-I in Patients with Thalassemia Major and Other Haematological Disorders*. Nature **201**: 1138-1139.
- Gilboa-Garber, N., Avichezer, D. y Garber, N.C. (1997). Bacterial Lectins: Properties, Structure, Effects, Function and Applications. *Glycosciences: Status and Perspectives*. H. J. Gabius and S. Gabius. Weinheim, Chapman & Hall: 369-398.
- Gilboa-Garber, N., Katcoff, D.J. y Garber, N.C. (2000). *Identification and Characterization of Pseudomonas Aeruginosa Pa-Iil Lectin Gene and Protein Compared to Pa-Il*. FEMS Immunol Med Microbiol **29**: 53-57.
- Giollant, M., Guillot, J., Damez, M., Dusser, M., Didier, P. y Didier, E. (1993). *Characterization of a Lectin from Lactarius Deterrimus (Research on the Possible Involvement of the Fungal Lectin in Recognition between Mushroom and Spruce During the Early Stages of Mycorrhizae Formation)*. Plant Physiol **101**: 513-522.
- Godoy, C.A., de las Rivas, B., Bezbradica, D., Bolívar, J.M., López-Gallego, F., Fernández-Lorente, G. y Guisán, J.M. (2011). *Reactivation of a Thermostable Lipase by Solid Phase Unfolding/Refolding Effect of Cysteine Residues on Refolding Efficiency*. Enzyme Microb. Technol. **49**: 388-394.
- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. y Oliver, S.G. (1996). *Life with 6000 Genes*. Science **274**: 546, 563-547.
- Gohler, A., Andre, S., Kaltner, H., Sauer, M., Gabius, H.J. y Doose, S. (2010). *Hydrodynamic Properties of Human Adhesion/Growth-Regulatory Galectins Studied by Fluorescence Correlation Spectroscopy*. Biophys J **98**: 3044-3053.
- Goldstein, I.J. y Poretz, R.D. (1986). Isolation, Physicochemical Characterization and Carbohydrate-Binding Specificity of Lectins. *The Lectins: Properties, Functions and Applications in Biology and Medicine*. I. E. Liener, N. Sharon and I. J. Goldstein. Orlando, FL, Academic Press: 35-247.
- Goldstein, I.J. y Winter, H.C. (2007). Mushroom Lectins. *Comprehensive Glycoscience: From Chemistry to Systems Biology*. J. P. Kamerling. Amsterdam, Elsevier Ltd: 601-621.
- Goloubinoff, P., Gatenby, A.A. y Lorimer, G.H. (1989). *GroE Heat-Shock Proteins Promote Assembly of Foreign Prokaryotic Ribulose Bisphosphate Carboxylase Oligomers in Escherichia Coli*. Nature **337**: 44-47.
- Gómez-Moreno, C. y Sancho, J. (2003). *Estructura De Proteínas*, Ariel Ciencia, Barcelona

González, B., Baños-Sanz, J.I., Villate, M., Brearley, C.A. y Sanz-Aparicio, J. (2010). *Inositol 1,3,4,5,6-Pentakisphosphate 2-Kinase Is a Distant Ipk Member with a Singular Inositide Binding Site for Axial 2-Oh Recognition*. Proc Natl Acad Sci U S A **107**: 9608-9613.

Gosavi, S., Whitford, P.C., Jennings, P.A. y Onuchic, J.N. (2008). *Extracting Function from a Beta-Trefoil Folding Motif*. Proc Natl Acad Sci U S A **105**: 10384-10389.

Grahn, E., Askarieh, G., Holmner, A., Tateno, H., Winter, H.C., Goldstein, I.J. y Krengel, U. (2007). *Crystal Structure of the Marasmius Oreades Mushroom Lectin in Complex with a Xenotransplantation Epitope*. J Mol Biol **369**: 710-721.

Grahn, E.M., Winter, H.C., Tateno, H., Goldstein, I.J. y Krengel, U. (2009). *Structural Characterization of a Lectin from the Mushroom Marasmius Oreades in Complex with the Blood Group B Trisaccharide and Calcium*. J Mol Biol **390**: 457-466.

Gruebele, M. (2005). *Fast Protein Folding: Evolution Meets Physics*. C R Biol **328**: 701-712.

Guillot, J., Genaud, L., Gueugnot, J. y Damez, M. (1983). *Purification and Properties of Two Hemagglutinins of the Mushroom Laccaria Amethystina*. Biochemistry **22**: 5365-5369.

Guillot, J., Giollant, M., Damez, M. y Dusser, M. (1991). *Isolation and Characterization of a Lectin from the Mushroom, Lactarius Deliciosus*. J Biochem **109**: 840-845.

Guillot, J. y Konska, G. (1997). *Lectins in Higher Fungi*. Biochem Syst Eco **25**: 203-230.

Gupta, G. y Surolia, A. (2012). Capítulo 1: Glycomics: An Overview of the Complex Glycocode. *Biochemical Roles of Eucaryotic Cell Surface Macromolecules*. P. R. Sudhakaran and A. Surolia. Nueva York, Springer: 1-14.

Hamshou, M., Smagghe, G., Shahidi-Noghabi, S., De Geyter, E., Lannoo, N. y Van Damme, E.J. (2010). *Insecticidal Properties of Sclerotinia Sclerotiorum Agglutinin and Its Interaction with Insect Tissues and Cells*. Insect Biochem Mol Biol **40**: 883-890.

Hamshou, M., Van Damme, E.J., Caccia, S., Cappelle, K., Vandenborre, G., Ghesquiere, B., Gevaert, K. y Smagghe, G. (2013). *High Entomotoxicity and Mechanism of the Fungal Galnac/Gal-Specific Rhizoctonia Solani Lectin in Pest Insects*. J Insect Physiol **59**: 295-305.

Hansen, J.E., Nielsen, C.M., Nielsen, C., Heegaard, P., Mathiesen, L.R. y Nielsen, J.O. (1989). *Correlation between Carbohydrate Structures on the Envelope Glycoprotein Gp120 of Hiv-1 and Hiv-2 and Syncytium Inhibition with Lectins*. AIDS **3**: 635-641.

Harada, J.J., Spadaro-Tank, J., Maxwell, J.C., Schnell, D.J. y Etzler, M.E. (1990). *Two Lectin Genes Differentially Expressed in Dolichos Biflorus Differ Primarily by a 116-Base Pair Sequence in Their 5' Flanking Regions*. J Biol Chem **265**: 4997-5001.

Harrison, M.J. (2005). *Signaling in the Arbuscular Mycorrhizal Symbiosis*. Annu Rev Microbiol **59**: 19-42.

Hartley, J.L. (2006). *Cloning Technologies for Protein Expression and Purification*. Curr Opin Biotechnol **17**: 359-366.

Hartz, D., Mcpheeters, D.S., Green, L. y Gold, L. (1991). *Detection of Escherichia Coli Ribosome Binding at Translation Initiation Sites in the Absence of Trna*. J Mol Biol **218**: 99-105.

Hassell, A.M., An, G., Bledsoe, R.K., Bynum, J.M., Carter, H.L., 3rd, Deng, S.J., Gampe, R.T., Grisard, T.E., Madauss, K.P., Nolte, R.T., Rocque, W.J., Wang, L., Weaver, K.L., Williams, S.P., Wisely, G.B., Xu, R. y Shewchuk, L.M. (2007). *Crystallization of Protein-Ligand Complexes*. Acta Crystallogr D Biol Crystallogr **63**: 72-79.

Hazes, B. (1996). *The (Q<sub>xw</sub>)<sub>3</sub> Domain: A Flexible Lectin Scaffold*. Protein Sci **5**: 1490-1501.

Heim, R., Cubitt, A.B. y Tsien, R.Y. (1995). *Improved Green Fluorescence*. Nature **373**: 663-664.

Helmy, M., Lombard, S. y Pieroni, G. (1999). *Ricin Rca60: Evidence of Its Phospholipase Activity*. Biochem Biophys Res Commun **258**: 252-255.

Henrissat, B. (1991). *A Classification of Glycosyl Hydrolases Based on Amino Acid Sequence Similarities*. Biochem J **280** ( Pt 2): 309-316.

Hermanson, G.T. (2008). *Bioconjugate Techniques*, Academic Press, San Diego

Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., Huhndorf, S., James, T., Kirk, P.M., Lücking, R., Thorsten Lumbsch, H., Lutzoni, F., Matheny, P.B., Mclaughlin, D.J., Powell, M.J., Redhead, S., Schoch, C.L., Spatafora, J.W., Stalpers, J.A., Vilgalys, R., Aime, M.C., Aptroot, A., Bauer, R., Begerow, D., Benny, G.L., Castlebury, L.A., Crous, P.W., Dai, Y.C., Gams, W., Geiser, D.M., Griffith, G.W., Gueidan, C., Hawksworth, D.L., Hestmark, G., Hosaka, K., Humber, R.A., Hyde, K.D., Ironside, J.E., Koljalg, U., Kurtzman, C.P., Larsson, K.H., Lichtwardt, R., Longcore, J., Miadlikowska, J., Miller, A., Moncalvo, J.M., Mozley-Standridge, S., Oberwinkler, F., Parmasto, E., Reeb, V., Rogers, J.D., Roux, C., Ryvarde, L., Sampaio, J.P., Schussler, A., Sugiyama, J., Thorn, R.G., Tibell, L., Untereiner, W.A., Walker, C., Wang, Z., Weir, A., Weiss, M., White, M.M., Winka, K., Yao, Y.J. y Zhang, N. (2007). *A Higher-Level Phylogenetic Classification of the Fungi*. Mycol Res **111**: 509-547.

Hillenkamp, F., Karas, M., Beavis, R.C. y Chait, B.T. (1991). *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Biopolymers*. Anal Chem **63**: 1193A-1203A.

Hirabayashi, J., Yamada, M., Kuno, A. y Tateno, H. (2013). *Lectin Microarrays: Concept, Principle and Applications*. Chem Soc Rev **42**: 4443-4458.

Horejsi, V. y Kocourek, J. (1978). *Studies on Lectins. Xxxvi. Properties of Some Lectins Prepared by Affinity Chromatography on O-Glycosyl Polyacrylamide Gels*. Biochim Biophys Acta **538**: 299-215.

Houseman, B.T. y Mrksich, M. (2002). *Carbohydrate Arrays for the Evaluation of Protein Binding and Enzymatic Modification*. Chem Biol **9**: 443-454.

Houser, J., Komarek, J., Kostlanova, N., Cioci, G., Imberty, A. y Wimmerova, M. (2012). *Protein Oligomerization in Aleuria Aurantia Lectin Family - Importance and Difficulties*. Mater Struct Chem Biol Phys Technol **12**: 20-21.

Houzelstein, D., Goncalves, I.R., Fadden, A.J., Sidhu, S.S., Cooper, D.N., Drickamer, K., Leffler, H. y Poirier, F. (2004). *Phylogenetic Analysis of the Vertebrate Galectin Family*. Mol Biol Evol **21**: 1177-1187.

Hu, S. y Wong, D.T. (2009). *Lectin Microarray*. Proteomics Clin Appl **3**: 148-154.

Imberty, A., Mitchell, E.P. y Wimmerova, M. (2005). *Structural Basis of High-Affinity Glycan Recognition by Bacterial and Fungal Lectins*. Curr Opin Struct Biol **15**: 525-534.

Jablonski, A. (1957). *Decay of Photoluminescence of Solutions* Acta Physica Polonica **16**: 471-479.

Jager, M., Zhang, Y., Bieschke, J., Nguyen, H., Dendle, M., Bowman, M.E., Noel, J.P., Gruebele, M. y Kelly, J.W. (2006). *Structure-Function-Folding Relationship in a Ww Domain*. Proc Natl Acad Sci U S A **103**: 10648-10653.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., Lumbsch, H.T., Rauhut, A., Reeb, V., Arnold, A.E., Amtoft, A., Stajich, J.E., Hosaka, K., Sung, G.H., Johnson, D., O'rourke, B., Crockett, M., Binder, M., Curtis, J.M., Slot, J.C., Wang, Z., Wilson, A.W., Schussler, A., Longcore, J.E., O'donnell, K., Mozley-Standridge, S., Porter, D., Letcher, P.M., Powell, M.J., Taylor, J.W., White, M.M., Griffith, G.W., Davies, D.R., Humber, R.A., Morton, J.B., Sugiyama, J., Rossman, A.Y., Rogers, J.D., Pfister, D.H., Hewitt, D., Hansen, K., Hambleton, S., Shoemaker, R.A., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Spotts, R.A., Serdani, M., Crous, P.W., Hughes, K.W., Matsuura, K., Langer, E., Langer, G., Untereiner, W.A., Lucking, R., Budel, B., Geiser, D.M., Aptroot, A., Diederich, P., Schmitt, I., Schultz, M., Yahr, R., Hibbett, D.S., Lutzoni, F., Mclaughlin, D.J., Spatafora, J.W. y Vilgalys, R. (2006). *Reconstructing the Early Evolution of Fungi Using a Six-Gene Phylogeny*. Nature **443**: 818-822.

Jancarik, J. y Kim, S.H. (1991). *Sparse-Matrix Sampling: A Screening Method for Crystallization of Proteins*. J Appl Crystallogr **24**: 409-411.

Johnson, C.M. (2013). *Differential Scanning Calorimetry as a Tool for Protein Folding and Stability*. Arch Biochem Biophys **531**: 100-109.

Jokerst, J.V., Chou, J., Camp, J.P., Wong, J., Lennart, A., Pollard, A.A., Floriano, P.N., Christodoulides, N., Simmons, G.W., Zhou, Y., Ali, M.F. y Mcdevitt, J.T. (2011). *Location of Biomarkers and Reagents within Agarose Beads of a Programmable Bio-Nano-Chip*. Small **7**: 613-624.

Kabsch, W. (2010). *Xds*. Acta Crystallogr D Biol Crystallogr **66**: 125-132.

Kadirvelraj, R., Grant, O.C., Goldstein, I.J., Winter, H.C., Tateno, H., Fadda, E. y Woods, R.J. (2011). *Structure and Binding Analysis of Polyporus Squamosus Lectin in Complex with the Neu5ac{Alpha}2-6gal{Beta}1-4glnac Human-Type Influenza Receptor*. Glycobiology **21**: 973-984.

Kahn, P.C. (1979). *The Interpretation of near-Ultraviolet Circular Dichroism*. Methods Enzymol **61**: 339-378.

Kalsi, G. y Etzler, M.E. (2000). *Localization of a Nod Factor-Binding Protein in Legume Roots and Factors Influencing Its Distribution and Expression*. Plant Physiol **124**: 1039-1048.



- Kaneko, T., Oguri, S., Kato, S. y Nagata, Y. (1993). *Developmental Appearance of Lectin During Fruit Body Formation in Pleurotus Cornucopiae*. J Gen Applied Microbiol **39**: 83-90.
- Kantardjieff, K.A. y Rupp, B. (2003). *Matthews Coefficient Probabilities: Improved Estimates for Unit Cell Contents of Proteins, DNA, and Protein-Nucleic Acid Complex Crystals*. Protein Sci **12**: 1865-1871.
- Karplus, P.A. y Diederichs, K. (2012). *Linking Crystallographic Model and Data Quality*. Science **336**: 1030-1033.
- Kataeva, I., Chang, J., Xu, H., Luan, C.H., Zhou, J., Uversky, V.N., Lin, D., Horanyi, P., Liu, Z.J., Ljungdahl, L.G., Rose, J., Luo, M. y Wang, B.C. (2005). *Improving Solubility of Shewanella Oneidensis Mr-1 and Clostridium Thermocellum Jw-20 Proteins Expressed into Escherichia Coli*. J Proteome Res **4**: 1942-1951.
- Katayama, H., McGill, M., Kearns, A., Brzozowski, M., Degner, N., Harnett, B., Kornilayev, B., Matkovic-Calogovic, D., Holyoak, T., Calvet, J.P., Gogol, E.P., Seed, J. y Fisher, M.T. (2009). *Strategies for Folding of Affinity Tagged Proteins Using Groel and Osmolytes*. J Struct Funct Genomics **10**: 57-66.
- Kellens, J.T.C., Goldstein, I.J. y Peumans, W.J. (1992). *Lectins in Different Members of the Sclerotiniaceae*. Mycol Res **96**: 495-502.
- Kelly, S.M., Jess, T.J. y Price, N.C. (2005). *How to Study Proteins by Circular Dichroism*. Biochim Biophys Acta **1751**: 119-139.
- Khan, F., Ahmad, A. y Khan, M.I. (2007). *Purification and Characterization of a Lectin from Endophytic Fungus Fusarium Solani Having Complex Sugar Specificity*. Arch Biochem Biophys **457**: 243-251.
- Khan, F. y Khan, M.I. (2011). *Fungal Lectins: Current Molecular and Biochemical Perspectives*. Int. J. Biol. Chem **5**: 1-20.
- Kim, C.W., Han, K.S., Ryu, K.S., Kim, B.H., Kim, K.H., Choi, S.I. y Seong, B.L. (2007). *N-Terminal Domains of Native Multidomain Proteins Have the Potential to Assist De Novo Folding of Their Downstream Domains in Vivo by Acting as Solubility Enhancers*. Protein Sci **16**: 635-643.
- Kim, D.E., Chivian, D. y Baker, D. (2004). *Protein Structure Prediction and Analysis Using the Robetta Server*. Nucleic Acids Res **32**: W526-531.
- Kingsman, S.M., Kingsman, A.J., Dobson, M.J., Mellor, J. y Roberts, N.A. (1985). *Heterologous Gene Expression in Saccharomyces Cerevisiae*. Biotechnol Genet Eng Rev **3**: 377-416.
- Kolaj, O., Spada, S., Robin, S. y Wall, J.G. (2009). *Use of Folding Modulators to Improve Heterologous Protein Production in Escherichia Coli*. Microb Cell Fact **8**: 9.
- Konarska, M.M., Grabowski, P.J., Padgett, R.A. y Sharp, P.A. (1985). *Characterization of the Branch Site in Lariat Rnas Produced by Splicing of Mrna Precursors*. Nature **313**: 552-557.

- Konska, G., Guillot, J., Dusser, M., Damez, M. y Botton, B. (1994). *Isolation and Characterization of an N-Acetyllactosamine-Binding Lectin from the Mushroom Laetiporus Sulfureus*. J Biochem **116**: 519-523.
- Korf, U., Kohl, T., Van Der Zandt, H., Zahn, R., Schleegeer, S., Ueberle, B., Wandschneider, S., Bechtel, S., Schnolzer, M., Otteleben, H., Wiemann, S. y Poustka, A. (2005). *Large-Scale Protein Expression for Proteome Research*. Proteomics **5**: 3571-3580.
- Kornfeld, K., Reitman, M.L. y Kornfeld, R. (1981). *The Carbohydrate-Binding Specificity of Pea and Lentil Lectins. Fucose Is an Important Determinant*. J Biol Chem **256**: 6633-6640.
- Krissinel, E. y Henrick, K. (2007). *Inference of Macromolecular Assemblies from Crystalline State*. J Mol Biol **372**: 774-797.
- Künzler, M., Bleuler-Martínez, S., Butschi, A., Garbani, M., Luthy, P., Hengartner, M.O. y Aebe, M. (2010). *Biotoxicity Assays for Fruiting Body Lectins and Other Cytoplasmic Proteins*. Methods Enzymol **480**: 141-150.
- Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.W., Roe, B.A. y Murphy, J.W. (2004). *Introns and Splicing Elements of Five Diverse Fungi*. Eukaryot Cell **3**: 1088-1100.
- Kupper, C.E., Rosencrantz, R.R., Henssen, B., Pelantova, H., Thones, S., Drozdova, A., Kren, V. y Elling, L. (2012). *Chemo-Enzymatic Modification of Poly-N-Acetyllactosamine (Lacnac) Oligomers and N,N-Diacetyllactosamine (Lacdinac) Based on Galactose Oxidase Treatment*. Beilstein J Org Chem **8**: 712-725.
- Ladero, M., Santos, A. y García-Ochoa, F. (2001). *Diffusion and Chemical Reaction Rates with Nonuniform Enzyme Distribution: An Experimental Approach*. Biotechnol Bioeng **72**: 458-467.
- Laemmli, U.K. (1970). *Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4*. Nature **227**: 680-685.
- Laine, R.A. (1997). *The Information-Storing Potential of the Sugar Code. Glycosciences: Status and Perspectives*. H. J. Gabius and S. Gabius. Weinheim, Chapman & Hall: 1-14.
- Lam, S.K. y Ng, T.B. (2011). *Lectins: Production and Practical Applications*. Appl Microbiol Biotechnol **89**: 45-55.
- Lannoo, N., Peumans, W.J., Pamel, E.V., Alvarez, R., Xiong, T.C., Hause, G., Mazars, C. y Van Damme, E.J. (2006). *Localization and in Vitro Binding Studies Suggest That the Cytoplasmic/Nuclear Tobacco Lectin Can Interact in Situ with High-Mannose and Complex N-Glycans*. FEBS Lett **580**: 6329-6337.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., Mcgettigan, P.A., Mcwilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. y Higgins, D.G. (2007). *Clustal W and Clustal X Version 2.0*. Bioinformatics **23**: 2947-2948.
- Larsen, P.E., Trivedi, G., Sreedasyam, A., Lu, V., Podila, G.K. y Collart, F.R. (2010). *Using Deep Rna Sequencing for the Structural Annotation of the Laccaria Bicolor Mycorrhizal Transcriptome*. PLoS One **5**: e9780.

Laskowski, R.A., Moss, D.S. y Thornton, J.M. (1993). *Main-Chain Bond Lengths and Bond Angles in Protein Structures*. J Mol Biol **231**: 1049-1067.

Laue, T.M. (1997). *Advances in Sedimentation Velocity Analysis*. Biophys J **72**: 395-396.

Lavallie, E.R., Diblasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F. y McCoy, J.M. (1993). *A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the E. Coli Cytoplasm*. Biotechnology (N Y) **11**: 187-193.

Lebowitz, J., Lewis, M.S. y Schuck, P. (2002). *Modern Analytical Ultracentrifugation in Protein Science: A Tutorial Review*. Protein Sci **11**: 2067-2079.

Lee, J. y Blaber, M. (2011). *Experimental Support for the Evolution of Symmetric Protein Architecture from a Simple Peptide Motif*. Proc Natl Acad Sci U S A **108**: 126-130.

Lemieux, R.U. (1996). *How Water Provides the Impetus for Molecular Recognition in Aqueous Solution*. Acc Chem Res **29**: 373-380.

Leonidas, D.D., Swamy, B.M., Bhat, A.G., Inamdar, S.R., Kosmopoulou, M.N., Chrysina, E.D. y Oikonomakos, N.G. (2003). *Crystallization and Preliminary X-Ray Crystallographic Analysis of Sclerotium Rolfsii Lectin*. Acta Crystallogr D Biol Crystallogr **59**: 363-365.

Leslie, A.G. (2006). *The Integration of Macromolecular Diffraction Data*. Acta Crystallogr D Biol Crystallogr **62**: 48-57.

Levitt, M. y Perutz, M.F. (1988). *Aromatic Rings Act as Hydrogen Bond Acceptors*. J Mol Biol **201**: 751-754.

Li, K., Ouyang, H., Lu, Y., Liang, J., Wilson, I.B. y Jin, C. (2011). *Repression of N-Glycosylation Triggers the Unfolded Protein Response (Upr) and Overexpression of Cell Wall Protein and Chitin in Aspergillus Fumigatus*. Microbiology **157**: 1968-1979.

Lichtenstein, R.G. y Rabinovich, G.A. (2013). *Glycobiology of Cell Death: When Glycans and Lectins Govern Cell Fate*. Cell Death Differ **20**: 976-986.

Liener, I.E. (1986). *Nutritional Significance of Lectins in the Diet. The Lectins. Properties, Functions, and Applications in Biology and Medicine*. I. E. Liener, N. Sharon and I. J. Goldstein. San Diego, CA, Academic Press: 527-552.

Liu, B., Cheng, Y., Bian, H.J. y Bao, J.K. (2009). *Molecular Mechanisms of Polygonatum Cyrtonema Lectin-Induced Apoptosis and Autophagy in Cancer Cells*. Autophagy **5**: 253-255.

Liu, Q., Wang, H. y Ng, T.B. (2004). *Isolation and Characterization of a Novel Lectin from the Wild Mushroom Xerocomus Spadiceus*. Peptides **25**: 7-10.

- Liu, Y., Chirino, A.J., Misulovin, Z., Leteux, C., Feizi, T., Nussenzweig, M.C. y Bjorkman, P.J. (2000). *Crystal Structure of the Cysteine-Rich Domain of Mannose Receptor Complexed with a Sulfated Carbohydrate Ligand*. J Exp Med **191**: 1105-1116.
- Liu, Z.B., Hou, Y.F., Min, D., Di, G.H., Wu, J., Shen, Z.Z. y Shao, Z.M. (2009). *Pa-Msha Inhibits Proliferation and Induces Apoptosis through the up-Regulation and Activation of Caspases in the Human Breast Cancer Cell Lines*. J Cell Biochem **108**: 195-206.
- Lobsanov, Y.D., Gitt, M.A., Leffler, H., Barondes, S.H. y Rini, J.M. (1993). *X-Ray Crystal Structure of the Human Dimeric S-Lac Lectin, L-14-Ii, in Complex with Lactose at 2.9-Å Resolution*. J Biol Chem **268**: 27034-27038.
- Logan, T.C., Clark, D.S., Stachowiak, T.B., Svec, F. y Frechet, J.M. (2007). *Photopatterning Enzymes on Polymer Monoliths in Microfluidic Devices for Steady-State Kinetic Analysis and Spatially Separated Multi-Enzyme Reactions*. Anal Chem **79**: 6592-6598.
- López-Gallego, F., Acebrón, I., Mancheño, J.M., Raja, S., Lillo, M.P. y Guisán Seijas, J.M. (2012). *Directed, Strong, and Reversible Immobilization of Proteins Tagged with a Beta-Trefoil Lectin Domain: A Simple Method to Immobilize Biomolecules on Plain Agarose Matrixes*. Bioconjug. Chem. **23**: 565-573.
- López-Lucendo, M.F., Solís, D., Andre, S., Hirabayashi, J., Kasai, K., Kaltner, H., Gabius, H.J. y Romero, A. (2004). *Growth-Regulatory Human Galectin-1: Crystallographic Characterisation of the Structural Changes Induced by Single-Site Mutations and Their Impact on the Thermodynamics of Ligand Binding*. J Mol Biol **343**: 957-970.
- Loris, R. (2002). *Principles of Structures of Animal and Plant Lectins*. Biochim Biophys Acta **1572**: 198-208.
- Lu, J., Teh, C., Kishore, U. y Reid, K.B. (2002). *Collectins and Ficolins: Sugar Pattern Recognition Molecules of the Mammalian Innate Immune System*. Biochim Biophys Acta **1572**: 387-400.
- Lu, Q., Bauer, J.C. y Greener, A. (1997). *Using Schizosaccharomyces Pombe as a Host for Expression and Purification of Eukaryotic Proteins*. Gene **200**: 135-144.
- Luft, J.R., Wolfley, J.R. y Snell, E.H. (2011). *What's in a Drop? Correlating Observations and Outcomes to Guide Macromolecular Crystallization Experiments*. Cryst Growth Des **11**: 651-663.
- Lutzoni, F., Kauff, F., Cox, C.J., Mclaughlin, D., Celio, G., Dentinger, B., Padamsee, M., Hibbett, D., James, T.Y., Baloch, E., Grube, M., Reeb, V., Hofstetter, V., Schoch, C., Arnold, A.E., Miadlikowska, J., Spatafora, J., Johnson, D., Hambleton, S., Crockett, M., Shoemaker, R., Sung, G.H., Lucking, R., Lumbsch, T., O'donnell, K., Binder, M., Diederich, P., Ertz, D., Gueidan, C., Hansen, K., Harris, R.C., Hosaka, K., Lim, Y.W., Matheny, B., Nishida, H., Pfister, D., Rogers, J., Rossman, A., Schmitt, I., Sipman, H., Stone, J., Sugiyama, J., Yahr, R. y Vilgalys, R. (2004). *Assembling the Fungal Tree of Life: Progress, Classification, and Evolution of Subcellular Traits*. Am J Bot **91**: 1446-1480.
- Malhotra, A. (2009). *Tagging for Protein Expression*. Methods Enzymol **463**: 239-258.
- Mancheño, J.M., Tateno, H., Goldstein, I.J., Martínez-Ripoll, M. y Hermoso, J.A. (2005). *Structural Analysis of the Laetiporus Sulphureus Hemolytic Pore-Forming Lectin in Complex with Sugars*. J Biol Chem **280**: 17251-17259.

Marblestone, J.G., Edavettal, S.C., Lim, Y., Lim, P., Zuo, X. y Butt, T.R. (2006). *Comparison of Sumo Fusion Technology with Traditional Gene Fusion Systems: Enhanced Expression and Solubility with Sumo*. Protein Sci **15**: 182-189.

Marín-Navarro, J., Talens-Perales, D., Oude-Vrielink, A., Canada, F.J. y Polaina, J. (2014). *Immobilization of Thermostable Beta-Galactosidase on Epoxy Support and Its Use for Lactose Hydrolysis and Galactooligosaccharides Biosynthesis*. World J Microbiol Biotechnol **30**: 989-998.

Marradi, M., Chiodo, F., Garcia, I. y Penades, S. (2013). *Glyconanoparticles as Multifunctional and Multimodal Carbohydrate Systems*. Chem Soc Rev **42**: 4728-4745.

Martin, F., Aerts, A., Ahren, D., Brun, A., Danchin, E.G., Duchaussoy, F., Gibon, J., Kohler, A., Lindquist, E., Pereda, V., Salamov, A., Shapiro, H.J., Wuyts, J., Blaudez, D., Buee, M., Brokstein, P., Canback, B., Cohen, D., Courty, P.E., Coutinho, P.M., Delaruelle, C., Detter, J.C., Deveau, A., Difazio, S., Duplessis, S., Fraissinet-Tachet, L., Lucic, E., Frey-Klett, P., Fourrey, C., Feussner, I., Gay, G., Grimwood, J., Hoegger, P.J., Jain, P., Kilaru, S., Labbe, J., Lin, Y.C., Legue, V., Le Tacon, F., Marmeisse, R., Melayah, D., Montanini, B., Muratet, M., Nehls, U., Niculita-Hirzel, H., Oudot-Le Secq, M.P., Peter, M., Quesneville, H., Rajashekar, B., Reich, M., Rouhier, N., Schmutz, J., Yin, T., Chalot, M., Henrissat, B., Kues, U., Lucas, S., Van De Peer, Y., Podila, G.K., Polle, A., Pukkila, P.J., Richardson, P.M., Rouze, P., Sanders, I.R., Stajich, J.E., Tunlid, A., Tuskan, G. y Grigoriev, I.V. (2008). *The Genome of Laccaria Bicolor Provides Insights into Mycorrhizal Symbiosis*. Nature **452**: 88-92.

Martin, F. y Nehls, U. (2009). *Harnessing Ectomycorrhizal Genomics for Ecological Insights*. Curr Opin Plant Biol **12**: 508-515.

Martin, F. y Selosse, M.A. (2008). *The Laccaria Genome: A Symbiont Blueprint Decoded*. New Phytol **180**: 296-310.

Mattos, C. (2002). *Protein-Water Interactions in a Dynamic World*. Trends Biochem Sci **27**: 203-208.

Mayhew, M., Da Silva, A.C., Martin, J., Erdjument-Bromage, H., Tempst, P. y Hartl, F.U. (1996). *Protein Folding in the Central Cavity of the Groel-Groes Chaperonin Complex*. Nature **379**: 420-426.

Mccormick, M. y Berg, J. (1997). *Purification and S-Tag Detection of Cbd Fusion Proteins*. inNovations **7**: 12-15.

Mccoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. y Read, R.J. (2007). *Phaser Crystallographic Software*. J Appl Crystallogr **40**: 658-674.

Mcnamara, P.J., Cuevas, W.A. y Songer, J.G. (1995). *Toxic Phospholipases D of Corynebacterium Pseudotuberculosis, C. Ulcerans and Arcanobacterium Haemolyticum: Cloning and Sequence Homology*. Gene **156**: 113-118.

Mcperson, A. (2004). *Introduction to Protein Crystallization*. Methods **34**: 254-265.

Melton, J.A., Parker, M.W., Rossjohn, J., Buckley, J.T. y Tweten, R.K. (2004). *The Identification and Structure of the Membrane-Spanning Domain of the Clostridium Septicum Alpha Toxin*. J Biol Chem **279**: 14315-14322.

- Merritt, E.A. (2012). *To B or Not to B: A Question of Resolution?* Acta Crystallogr D Biol Crystallogr **68**: 468-477.
- Mikeska, R., Wacker, R., Arni, R., Singh, T.P., Mikhailov, A., Gabdoulkhakov, A., Voelter, W. y Betzel, C. (2005). *Mistletoe Lectin I in Complex with Galactose and Lactose Reveals Distinct Sugar-Binding Properties*. Acta Crystallogr Sect F Struct Biol Cryst Commun **61**: 17-25.
- Mitchell, S.W. (1860). *Researches Upon the Venom of the Rattlesnake*. Smithson Contrib Knowl **XII**: 89-90.
- Montenegro, J.M., Grazu, V., Sukhanova, A., Agarwal, S., De La Fuente, J.M., Nabiev, I., Greiner, A. y Parak, W.J. (2013). *Controlled Antibody/(Bio-) Conjugation of Inorganic Nanoparticles for Targeted Delivery*. Adv. Drug. Deliv. Rev. **65**: 677-688.
- Montfort, W., Villafranca, J.E., Monzingo, A.F., Ernst, S.R., Katzin, B., Rutenber, E., Xuong, N.H., Hamlin, R. y Robertus, J.D. (1987). *The Three-Dimensional Structure of Ricin at 2.8 Å*. J Biol Chem **262**: 5398-5403.
- Moros, M., Pelaz, B., Lopez-Larrubia, P., Garcia-Martin, M.L., Grazu, V. y De La Fuente, J.M. (2010). *Engineering Biofunctional Magnetic Nanoparticles for Biotechnological Applications*. Nanoscale **2**: 1746-1755.
- Mukhopadhyay, D. (2000). *The Molecular Evolutionary History of a Winged Bean Alpha-Chymotrypsin Inhibitor and Modeling of Its Mutations through Structural Analyses*. J Mol Evol **50**: 214-223.
- Murzin, A.G., Lesk, A.M. y Chothia, C. (1992). *Beta-Trefoil Fold. Patterns of Structure and Sequence in the Kunitz Inhibitors Interleukins-1 Beta and 1 Alpha and Fibroblast Growth Factors*. J Mol Biol **223**: 531-543.
- Nadig, G., Van Zant, L.C., Dixon, S.L. y Merz, K.M. (1998). *Charge-Transfer Interactions in Macromolecular Systems: A New View of the Protein/Water Interface*. J Am Chem Soc **120**: 5593-5594.
- Nallamsetty, S. y Waugh, D.S. (2006). *Solubility-Enhancing Proteins Mbp and Nusa Play a Passive Role in the Folding of Their Fusion Partners*. Protein Expr Purif **45**: 175-182.
- Nam, K.H., Kim, S.J., Kim, M.Y., Kim, J.H., Yeo, Y.S., Lee, C.M., Jun, H.K. y Hwang, K.Y. (2008). *Crystal Structure of Engineered Beta-Glucosidase from a Soil Metagenome*. Proteins **73**: 788-793.
- Notenboom, V., Boraston, A.B., Williams, S.J., Kilburn, D.G. y Rose, D.R. (2002). *High-Resolution Crystal Structures of the Lectin-Like Xylan Binding Domain from Streptomyces Lividans Xylanase 10a with Bound Substrates Reveal a Novel Mode of Xylan Binding*. Biochemistry **41**: 4246-4254.
- Nyame, A.K., Debose-Boyd, R., Long, T.D., Tsang, V.C. y Cummings, R.D. (1998). *Expression of Lex Antigen in Schistosoma Japonicum and S.Haematobium and Immune Responses to Lex in Infected Animals: Lack of Lex Expression in Other Trematodes and Nematodes*. Glycobiology **8**: 615-624.
- Nyame, A.K., Yoshino, T.P. y Cummings, R.D. (2002). *Differential Expression of Lacdinac, Fucosylated Lacdinac, and Lewis X Glycan Antigens in Intramolluscan Stages of Schistosoma Mansoni*. J Parasitol **88**: 890-897.

Nygren, P.A., Stahl, S. y Uhlen, M. (1994). *Engineering Proteins to Facilitate Bioprocessing*. Trends Biotechnol **12**: 184-188.

Ofek, I. y Sharon, N. (1988). *Lectinophagocytosis: A Molecular Mechanism of Recognition between Cell Surface Sugars and Lectins in the Phagocytosis of Bacteria*. Infect Immun **56**: 539-547.

Oliver, R.P. y Solomon, P.S. (2010). *New Developments in Pathogenicity and Virulence of Necrotrophs*. Curr Opin Plant Biol **13**: 415-419.

Olsen, E.B., Russell, J.B. y Henick-Kling, T. (1991). *Electrogenic L-Malate Transport by Lactobacillus Plantarum: A Basis for Energy Derivation from Malolactic Fermentation*. J Bacteriol **173**: 6199-6206.

Paddock, S.W. (1999). *Confocal Laser Scanning Microscopy*. Biotechniques **27**: 992-996, 998-1002, 1004.

Pandey, A., Selvakumar, P., Soccol, C.R. y Nigam, P. (1999). *Solid State Fermentation for the Production of Industrial Enzymes*. Curr Sci **77**: 149-162.

Pankhurst, Q.A., Connolly, J., Jones, S.K. y Dobson, J. (2003). *Applications of Magnetic Nanoparticles in Biomedicine*. Journal of Physics D, Applied Physics **36**: 167-181.

Park, S., Gildersleeve, J.C., Blixt, O. y Shin, I. (2013). *Carbohydrate Microarrays*. Chem Soc Rev **42**: 4310-4326.

Park, S. y Shin, I. (2002). *Fabrication of Carbohydrate Chips for Studying Protein-Carbohydrate Interactions*. Angew Chem Int Ed Engl **41**: 3180-3182.

Park, Y.W. (2009). *Bioactive Components in Milk and Dairy Products*, Wiley-Blackwell, Singapur

Paulson, J.C., Blixt, O. y Collins, B.E. (2006). *Sweet Spots in Functional Glycomics*. Nat Chem Biol **2**: 238-248.

Pel, H.J., De Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., Turner, G., De Vries, R.P., Albang, R., Albermann, K., Andersen, M.R., Bendtsen, J.D., Benen, J.A., Van Den Berg, M., Breestraat, S., Caddick, M.X., Contreras, R., Cornell, M., Coutinho, P.M., Danchin, E.G., Debets, A.J., Dekker, P., Van Dijk, P.W., Van Dijk, A., Dijkhuizen, L., Driessen, A.J., D'enfert, C., Geysens, S., Goosen, C., Groot, G.S., De Groot, P.W., Guillemette, T., Henrissat, B., Herweijer, M., Van Den Hombergh, J.P., Van Den Hondel, C.A., Van Der Heijden, R.T., Van Der Kaaij, R.M., Klis, F.M., Kools, H.J., Kubicek, C.P., Van Kuyk, P.A., Lauber, J., Lu, X., Van Der Maarel, M.J., Meulenberg, R., Menke, H., Mortimer, M.A., Nielsen, J., Oliver, S.G., Olsthoorn, M., Pal, K., Van Peij, N.N., Ram, A.F., Rinas, U., Roubos, J.A., Sagt, C.M., Schmoll, M., Sun, J., Ussery, D., Varga, J., Vervecken, W., Van De Vondervoort, P.J., Wedler, H., Wosten, H.A., Zeng, A.P., Van Ooyen, A.J., Visser, J. y Stam, H. (2007). *Genome Sequencing and Analysis of the Versatile Cell Factory Aspergillus Niger Cbs 513.88*. Nat Biotechnol **25**: 221-231.

Petukhov, M., Cregut, D., Soares, C.M. y Serrano, L. (1999). *Local Water Bridges and Protein Conformational Stability*. Protein Sci **8**: 1982-1989.

Peumans, W.J. y Van Damme, E.J. (1995). *The Role of Lectins in Plant Defence*. Histochem J **27**: 253-271.

- Piller, F., Cartron, J.P., Maranduba, A., Veyrieres, A., Leroy, Y. y Fournet, B. (1984). *Biosynthesis of Blood Group I Antigens. Identification of a Udp-Glcna:Glcna Beta 1-3gal(-R) Beta 1-6(Glcna to Gal) N-Acetylglucosaminyltransferase in Hog Gastric Mucosa*. J Biol Chem **259**: 13385-13390.
- Pohleven, J., Renko, M., Magister, S., Smith, D.F., Kunzler, M., Strukelj, B., Turk, D., Kos, J. y Sabotic, J. (2012). *Bivalent Carbohydrate Binding Is Required for Biological Activity of Clitocybe Nebularis Lectin (Cnl), the N,N'-Diacetyllactosediamine (Galnacbeta1-4glcna, Lacdinac)-Specific Lectin from Basidiomycete C. Nebularis*. J Biol Chem **287**: 10602-10612.
- Ponting, C.P. y Russell, R.B. (2000). *Identification of Distant Homologues of Fibroblast Growth Factors Suggests a Common Ancestor for All Beta-Trefoil Proteins*. J Mol Biol **302**: 1041-1047.
- Popov, A.N. y Bourenkov, G.P. (2003). *Choice of Data-Collection Parameters Based on Statistic Modelling*. Acta Crystallogr D Biol Crystallogr **59**: 1145-1153.
- Pryor, K.D. y Leiting, B. (1997). *High-Level Expression of Soluble Protein in Escherichia Coli Using a His6-Tag and Maltose-Binding-Protein Double-Affinity Fusion System*. Protein Expr Purif **10**: 309-319.
- Puchades, M., Nilsson, C.L., Emmett, M.R., Aldape, K.D., Ji, Y., Lang, F.F., Liu, T.J. y Conrad, C.A. (2007). *Proteomic Investigation of Glioblastoma Cell Lines Treated with Wild-Type P53 and Cytotoxic Chemotherapy Demonstrates an Association between Galectin-1 and P53 Expression*. J Proteome Res **6**: 869-875.
- Punta, M., Coghill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A. y Finn, R.D. (2012). *The Pfam Protein Families Database*. Nucleic Acids Res **40**: D290-301.
- Raes, J., Harrington, E.D., Singh, A.H. y Bork, P. (2007). *Protein Function Space: Viewing the Limits or Limited by Our View?* Curr Opin Struct Biol **17**: 362-369.
- Rahaie, M. y Kazemi, S.S. (2010). *Lectin-Based Biosensors: As Powerful Tools in Bioanalytical Applications* Biotechnology **9**: 428-443.
- Ramachandran, G.N., Ramakrishnan, C. y Sasisekharan, V. (1963). *Stereochemistry of Polypeptide Chain Configurations*. J Mol Biol **7**: 95-99.
- Raschke, T.M. (2006). *Water Structure and Interactions with Protein Surfaces*. Curr Opin Struct Biol **16**: 152-159.
- Rhodes, G. (2006). *Crystallography Made Crystal Clear*, Academic Press, Londres
- Rini, J.M. (1995). *Lectin Structure*. Annu Rev Biophys Biomol Struct **24**: 551-577.
- Rini, J.M. y Lobsanov, Y.D. (1999). *New Animal Lectin Structures*. Curr Opin Struct Biol **9**: 578-584.



- Rivas, G., Stafford, W. y Minton, A.P. (1999). *Characterization of Heterologous Protein-Protein Interactions Using Analytical Ultracentrifugation*. *Methods* **19**: 194-212.
- Rodríguez, H., Angulo, I., De Las Rivas, B., Campillo, N., Paez, J.A., Muñoz, R. y Mancheño, J.M. (2010). *P-Coumaric Acid Decarboxylase from Lactobacillus Plantarum: Structural Insights into the Active Site and Decarboxylation Catalytic Mechanism*. *Proteins* **78**: 1662-1676.
- Rodríguez-Arango, E., Arango, R., Adar, R., Galili, G. y Sharon, N. (1992). *Cloning, Sequence Analysis and Expression in Escherichia Coli of the Cdna Encoding a Precursor of Peanut Agglutinin*. *FEBS Lett* **307**: 185-189.
- Rohlf, M., Albert, M., Keller, N.P. y Kempken, F. (2007). *Secondary Chemicals Protect Mould from Fungivory*. *Biol Lett* **3**: 523-525.
- Roldos, V., Canada, F.J. y Jimenez-Barbero, J. (2011). *Carbohydrate-Protein Interactions: A 3d View by Nmr*. *Chembiochem* **12**: 990-1005.
- Roseman, S. (2001). *Reflections on Glycobiology*. *J Biol Chem* **276**: 41527-41542.
- Rossi, L.M., Quach, A.D. y Rosenzweig, Z. (2004). *Glucose Oxidase-Magnetite Nanoparticle Bioconjugate for Glucose Sensing*. *Anal Bioanal Chem* **380**: 606-613.
- Rubinson, K.A., Ladner, J.E., Tordova, M. y Gilliland, G.L. (2000). *Cryosalts: Suppression of Ice Formation in Macromolecular Crystallography*. *Acta Crystallogr D Biol Crystallogr* **56**: 996-1001.
- Rüdiger, H. (1998). *Plant Lectins - More Than Just Tools for Glycoscientists: Occurrence, Structure, and Possible Functions of Plant Lectins*. *Acta Anat (Basel)* **161**: 130-152.
- Rüdiger, H. y Gabius, H.J. (2001). *Plant Lectins: Occurrence, Biochemistry, Functions and Applications*. *Glycoconj J* **18**: 589-613.
- Russo Krauss, I., Merlino, A., Vergara, A. y Sica, F. (2013). *An Overview of Biological Macromolecule Crystallization*. *Int J Mol Sci* **14**: 11643-11691.
- Rutenber, E., Katzin, B.J., Ernst, S., Collins, E.J., Mlsna, D., Ready, M.P. y Robertus, J.D. (1991). *Crystallographic Refinement of Ricin to 2.5 Å*. *Proteins* **10**: 240-250.
- Sa, R.A., Santos, N.D., Da Silva, C.S., Napoleao, T.H., Gomes, F.S., Cavada, B.S., Coelho, L.C., Navarro, D.M., Bieber, L.W. y Paiva, P.M. (2009). *Larvicidal Activity of Lectins from Myracrodruon Urundeuva on Aedes Aegypti*. *Comp Biochem Physiol C Toxicol Pharmacol* **149**: 300-306.
- Safdar, M., Hashemi, P. y Adeli, M. (2013). *One-Step Synthesis of Agarose Coated Magnetic Nanoparticles and Their Application in the Solid Phase Extraction of Pd(II) Using a New Magnetic Field Agitation Device*. *Anal Chim Acta* **774**: 44-50.
- Saleemuddin, M. (1999). *Bioaffinity Based Immobilization of Enzymes*. *Adv Biochem Eng Biotechnol* **64**: 203-226.

Salomonsson, E., Carlsson, M.C., Osla, V., Hendus-Altenburger, R., Kahl-Knutson, B., Oberg, C.T., Sundin, A., Nilsson, R., Nordberg-Karlsson, E., Nilsson, U.J., Karlsson, A., Rini, J.M. y Leffler, H. (2010). *Mutational Tuning of Galectin-3 Specificity and Biological Function*. J Biol Chem **285**: 35079-35091.

Sambrook, J. y Russell, D.W. (2001). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Nueva York

Santana, S.D., Pina, A.S. y Roque, A.C. (2012). *Immobilization of Enterokinase on Magnetic Supports for the Cleavage of Fusion Proteins*. J Biotechnol **161**: 378-382.

Sasaki, N., Yoshida, H., Fuwa, T.J., Kinoshita-Toyoda, A., Toyoda, H., Hirabayashi, Y., Ishida, H., Ueda, R. y Nishihara, S. (2007). *Drosophila Beta 1,4-N-Acetylgalactosaminyltransferase-a Synthesizes the Lacdinac Structures on Several Glycoproteins and Glycosphingolipids*. Biochem Biophys Res Commun **354**: 522-527.

Schatz, P.J. (1993). *Use of Peptide Libraries to Map the Substrate Specificity of a Peptide-Modifying Enzyme: A 13 Residue Consensus Peptide Specifies Biotinylation in Escherichia Coli*. Biotechnology (N Y) **11**: 1138-1143.

Schreiber, G., Buckle, A.M. y Fersht, A.R. (1994). *Stability and Function: Two Constraints in the Evolution of Barstar and Other Proteins*. Structure **2**: 945-951.

Schubert, M., Bleuler-Martinez, S., Butschi, A., Walti, M.A., Egloff, P., Stutz, K., Yan, S., Collot, M., Mallet, J.M., Wilson, I.B., Hengartner, M.O., Aebi, M., Allain, F.H. y Kunzler, M. (2012). *Plasticity of the Beta-Trefoil Protein Fold in the Recognition and Control of Invertebrate Predators and Parasites by a Fungal Defence System*. PLoS Pathog **8**: e1002706.

Scott, N., Hatlelid, K.M., Mackenzie, N.E. y Carter, D.E. (1993). *Reactions of Arsenic(III) and Arsenic(V) Species with Glutathione*. Chem Res Toxicol **6**: 102-106.

Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S.H. y Rini, J.M. (1998). *X-Ray Crystal Structure of the Human Galectin-3 Carbohydrate Recognition Domain at 2.1-Å Resolution*. J Biol Chem **273**: 13047-13052.

Sethuraman, N. y Stadheim, T.A. (2006). *Challenges in Therapeutic Glycoprotein Production*. Curr Opin Biotechnol **17**: 341-346.

Seto, M., Ogawa, T., Kodama, K., Muramoto, K., Kanayama, Y., Sakai, Y., Chijiwa, T. y Ohno, M. (2008). *A Novel Recombinant System for Functional Expression of Myonecrotic Snake Phospholipase a(2) in Escherichia Coli Using a New Fusion Affinity Tag*. Protein Expr Purif **58**: 194-202.

Sharon, N. y Goldstein, I.J. (1998). *Lectins: More Than Insecticides*. Science **282**: 1049.

Sharon, N. y Lis, H. (2007). *Lectins*, Springer Netherlands, Dordrecht

Sheldon, R.A. y Van Pelt, S. (2013). *Enzyme Immobilisation in Biocatalysis: Why, What and How*. Chem Soc Rev **42**: 6223-6235.

- Shoichet, B.K., Baase, W.A., Kuroki, R. y Matthews, B.W. (1995). *A Relationship between Protein Stability and Protein Function*. Proc Natl Acad Sci U S A **92**: 452-456.
- Silva-Martín, N., Molina, R., Angulo, I., Mancheño, J.M., García, P. y Hermoso, J.A. (2010). *Crystallization and Preliminary Crystallographic Analysis of the Catalytic Module of Endolysin from Cp-7, a Phage Infecting Streptococcus Pneumoniae*. Acta Crystallogr Sect F Struct Biol Cryst Commun **66**: 670-673.
- Singh, R.S., Thakur, G. y Bhari, R. (2009). *Optimization of Culture Conditions and Characterization of a New Lectin from Aspergillus Niger*. Indian J Microbiol **49**: 219-222.
- Singh, R.S., Tiwary, A.K. y Bhari, R. (2008). *Screening of Aspergillus Species for Occurrence of Lectins and Their Characterization*. J Basic Microbiol **48**: 112-117.
- Skamnaki, V.T., Peumans, W.J., Kantsadi, A.L., Cubeta, M.A., Plas, K., Pakala, S., Zographos, S.E., Smagghe, G., Nierman, W.C., Van Damme, E.J. y Leonidas, D.D. (2013). *Structural Analysis of the Rhizoctonia Solani Agglutinin Reveals a Domain-Swapping Dimeric Assembly*. FEBS J **280**: 1750-1763.
- Smith, D.B. y Johnson, K.S. (1988). *Single-Step Purification of Polypeptides Expressed in Escherichia Coli as Fusions with Glutathione S-Transferase*. Gene **67**: 31-40.
- Smith, D.F., Song, X. y Cummings, R.D. (2010). *Use of Glycan Microarrays to Explore Specificity of Glycan-Binding Proteins*. Methods Enzymol **480**: 417-444.
- Smith, S.E. y Read, D.J. (2008). *Mycorrhizal Symbiosis*, Academic Press, Nueva York
- Spano, G., Rinaldi, A., Ugliano, M., Moio, L., Beneduce, L. y Massa, S. (2005). *A Beta-Glucosidase Gene Isolated from Wine Lactobacillus Plantarum Is Regulated by Abiotic Stresses*. J Appl Microbiol **98**: 855-861.
- Spencer, R.D. y Weber, G. (1970). *Influence of Brownian Rotations and Energy Transfer Upon the Measurements of Fluorescence Lifetime*. J Chem Phys **52**: 1654-1663.
- Spiro, R.G. (1960). *Studies on Fetuin, a Glycoprotein of Fetal Serum. I. Isolation, Chemical Composition, and Physicochemical Properties*. J Biol Chem **235**: 2860-2869.
- Stajich, J.E., Dietrich, F.S. y Roy, S.W. (2007). *Comparative Genomic Analysis of Fungal Genomes Reveals Intron-Rich Ancestors*. Genome Biol **8**: R223.
- Stanke, M. y Morgenstern, B. (2005). *Augustus: A Web Server for Gene Prediction in Eukaryotes That Allows User-Defined Constraints*. Nucleic Acids Res **33**: W465-467.
- Stillmark, H. (1888). Ueber Ricin, Ein Giftiges Ferment Aus Den Samen Von Ricinus Comm. L. Und Einigen Anderen Euphorbiaceen. Dorpat (ahora Tartu), Estonia, Universidad de Dorpat. MD Thesis.
- Stowell, S.R., Arthur, C.M., Dias-Baruffi, M., Rodrigues, L.C., Gourdine, J.P., Heimburg-Molinaro, J., Ju, T., Molinaro, R.J., Rivera-Marrero, C., Xia, B., Smith, D.F. y Cummings, R.D. (2010). *Innate Immune Lectins Kill Bacteria Expressing Blood Group Antigen*. Nat Med **16**: 295-301.

Strickland, E.H. (1974). *Aromatic Contributions to Circular Dichroism Spectra of Proteins*. CRC Crit Rev Biochem **2**: 113-175.

Structural Genomics, C., China Structural Genomics, C., Northeast Structural Genomics, C., Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., Dhe-Paganon, S., Park, H.W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S.H., Rao, Z., Shi, Y., Terwilliger, T.C., Kim, C.Y., Hung, L.W., Waldo, G.S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J.L., Stevens, R.C., Lesley, S.A., Wilson, I.A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M.I., Eschenfeldt, W.H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S.K., Emtage, J.S., Sauder, J.M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S.C., Bonanno, J.B., Fiser, A., Swaminathan, S., Studier, F.W., Chance, M.R., Sali, A., Acton, T.B., Xiao, R., Zhao, L., Ma, L.C., Hunt, J.F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C.K., Wang, D., Wang, H., Jiang, M., Montelione, G.T., Stuart, D.I., Owens, R.J., Daenke, S., Schutz, A., Heinemann, U., Yokoyama, S., Bussow, K. y Gunsalus, K.C. (2008). *Protein Production and Purification*. Nat Methods **5**: 135-146.

Suckau, D., Resemann, A., Schuerenberg, M., Hufnagel, P., Franzen, J. y Holle, A. (2003). *A Novel Maldi Lift-Tof/Tof Mass Spectrometer for Proteomics*. Anal Bioanal Chem **376**: 952-965.

Sulzenbacher, G., Roig-Zamboni, V., Peumans, W.J., Rouge, P., Van Damme, E.J. y Bourne, Y. (2010). *Crystal Structure of the Galnac/Gal-Specific Agglutinin from the Phytopathogenic Ascomycete Sclerotinia Sclerotiorum Reveals Novel Adaptation of a Beta-Trefoil Domain*. J Mol Biol **400**: 715-723.

Sun, C., Lee, J.S. y Zhang, M. (2008). *Magnetic Nanoparticles in Mr Imaging and Drug Delivery*. Adv. Drug Deliv. Rev. **60**: 1252-1265.

Sun, C., Lee, J.S. y Zhang, M. (2008). *Magnetic Nanoparticles in Mr Imaging and Drug Delivery*. Adv Drug Deliv Rev **60**: 1252-1265.

Suzuki, R., Kuno, A., Hasegawa, T., Hirabayashi, J., Kasai, K.I., Momma, M. y Fujimoto, Z. (2009). *Sugar-Complex Structures of the C-Half Domain of the Galactose-Binding Lectin Ew29 from the Earthworm Lumbricus Terrestris*. Acta Crystallogr D Biol Crystallogr **65**: 49-57.

Swaminathan, G.J., Leonidas, D.D., Savage, M.P., Ackerman, S.J. y Acharya, K.R. (1999). *Selective Recognition of Mannose by the Human Eosinophil Charcot-Leyden Crystal Protein (Galectin-10): A Crystallographic Study at 1.8 Å Resolution*. Biochemistry **38**: 15406.

Sweet, R.M., Wright, H.T., Janin, J., Chothia, C.H. y Blow, D.M. (1974). *Crystal Structure of the Complex of Porcine Trypsin with Soybean Trypsin Inhibitor (Kunitz) at 2.6-Å Resolution*. Biochemistry **13**: 4212-4228.

Tanaka, H., Chiba, H., Inokoshi, J., Kuno, A., Sugai, T., Takahashi, A., Ito, Y., Tsunoda, M., Suzuki, K., Takenaka, A., Sekiguchi, T., Umeyama, H., Hirabayashi, J. y Omura, S. (2009). *Mechanism by Which the Lectin Actinohivin Blocks Hiv Infection of Target Cells*. Proc Natl Acad Sci U S A **106**: 15633-15638.

Tang, J., Liu, Y., Yin, P., Yao, G., Yan, G., Deng, C. y Zhang, X. (2010). *Concanavalin α-Immobilized Magnetic Nanoparticles for Selective Enrichment of Glycoproteins and Application to Glycoproteomics in Hepatocellular Carcinoma Cell Line*. Proteomics **10**: 2000-2014.

- Tartaj, P., Morales, M.P., Veintemillas-Verdaguer, S., González-Carreño, T. y Serna, C.J. (2003). *The Preparation of Magnetic Nanoparticles for Applications in Biomedicine*. J. Phys. D: Appl. Phys. **36**: 182-197.
- Tateno, H. y Goldstein, I.J. (2003). *Molecular Cloning, Expression, and Characterization of Novel Hemolytic Lectins from the Mushroom Laetiporus Sulphureus, Which Show Homology to Bacterial Toxins*. J Biol Chem **278**: 40455-40463.
- Tateno, H., Kuno, A., Itakura, Y. y Hirabayashi, J. (2010). *A Versatile Technology for Cellular Glycomics Using Lectin Microarray*. Methods Enzymol **478**: 181-195.
- Tateno, H., Winter, H.C. y Goldstein, I.J. (2004). *Cloning, Expression in Escherichia Coli and Characterization of the Recombinant Neu5acalpha2,6galbeta1,4glcnac-Specific High-Affinity Lectin and Its Mutants from the Mushroom Polyporus Squamosus*. Biochem J **382**: 667-675.
- Tateno, Y., Ikeo, K., Imanishi, T., Watanabe, H., Endo, T., Yamaguchi, Y., Suzuki, Y., Takahashi, K., Tsunoyama, K., Kawai, M., Kawanishi, Y., Naitou, K. y Gojobori, T. (1997). *Evolutionary Motif and Its Biological and Structural Significance*. J Mol Evol **44 Suppl 1**: S38-43.
- Teeuwen, R.L.M., Zuilhof, H., De Wolf, F.A. y Van Hest, J.C.M. (2009). *Temperature-Controlled Positioning of Fusion Proteins in Microreactors*. Soft Matter. **5**: 2261-2268.
- Terpe, K. (2003). *Overview of Tag Protein Fusions: From Molecular and Biochemical Fundamentals to Commercial Systems*. Appl Microbiol Biotechnol **60**: 523-533.
- Tetala, K.K. y Van Beek, T.A. (2010). *Bioaffinity Chromatography on Monolithic Supports*. J Sep Sci **33**: 422-438.
- Thain, A., Gaston, K., Jenkins, O. y Clarke, A.R. (1996). *A Method for the Separation of Gst Fusion Proteins from Co-Purifying Groel*. Trends Genet **12**: 209-210.
- Thomas, J.G., Ayling, A. y Baneyx, F. (1997). *Molecular Chaperones, Folding Catalysts, and the Recovery of Active Recombinant Proteins from E. Coli. To Fold or to Refold*. Appl Biochem Biotechnol **66**: 197-238.
- Tielker, D., Rosenau, F., Bartels, K.M., Rosenbaum, T. y Jaeger, K.E. (2006). *Lectin-Based Affinity Tag for One-Step Protein Purification*. Biotechniques **41**: 327-332.
- Tomme, P., Boraston, A., Mclean, B., Kormos, J., Creagh, A.L., Sturch, K., Gilkes, N.R., Haynes, C.A., Warren, R.A. y Kilburn, D.G. (1998). *Characterization and Affinity Applications of Cellulose-Binding Domains*. J Chromatogr B Biomed Sci Appl **715**: 283-296.
- Toone, E.J. (1994). *Structure and Energetics of Protein-Carbohydrate Complexes*. Curr Opin Struct Biol **4**: 719-728.
- Townsend, R.R., Hilliker, E., Li, Y.T., Laine, R.A., Bell, W.R. y Lee, Y.C. (1982). *Carbohydrate Structure of Human Fibrinogen. Use of 300-Mhz 1h-Nmr to Characterize Glycosidase-Treated Glycopeptides*. J Biol Chem **257**: 9704-9710.

- Transue, T.R., Smith, A.K., Mo, H., Goldstein, I.J. y Saper, M.A. (1997). *Structure of Benzyl T-Antigen Disaccharide Bound to Amaranthus Caudatus Agglutinin*. *Nat Struct Biol* **4**: 779-783.
- Tremblay, R., Diao, H., Huner, N., Jevnikar, A.M. y Ma, S. (2011). *The Development, Characterization, and Demonstration of a Novel Strategy for Purification of Recombinant Proteins Expressed in Plants*. *Transgenic Res* **20**: 1357-1366.
- Trigueros, V., Lougarre, A., Ali-Ahmed, D., Rahbe, Y., Guillot, J., Chavant, L., Fournier, D. y Paquereau, L. (2003). *Xerocomus Chrysenteron Lectin: Identification of a New Pesticidal Protein*. *Biochim Biophys Acta* **1621**: 292-298.
- Trombetta, E.S. y Helenius, A. (1998). *Lectins as Chaperones in Glycoprotein Folding*. *Curr Opin Struct Biol* **8**: 587-592.
- Tronchin, G., Esnault, K., Sanchez, M., Larcher, G., Marot-Leblond, A. y Bouchara, J.P. (2002). *Purification and Partial Characterization of a 32-Kilodalton Sialic Acid-Specific Lectin from Aspergillus Fumigatus*. *Infect Immun* **70**: 6891-6895.
- Tschampel, S.M. y Woods, R.J. (2003). *Quantifying the Role of Water in Protein-Carbohydrate Interactions*. *J Phys Chem A* **107**: 9175-9181.
- Tsien, R.Y. (1998). *The Green Fluorescent Protein*. *Annu Rev Biochem* **67**: 509-544.
- Twu, Y.C., Hsieh, C.Y., Lin, M., Tzeng, C.H., Sun, C.F. y Yu, L.C. (2010). *Phosphorylation Status of Transcription Factor C/Ebpalpha Determines Cell-Surface Poly-Lacnac Branching (I Antigen) Formation in Erythropoiesis and Granulopoiesis*. *Blood* **115**: 2491-2499.
- Uskova, M.A., Borst, J.W., Hink, M.A., Van Hoek, A., Schots, A., Klyachko, N.L. y Visser, A.J. (2000). *Fluorescence Dynamics of Green Fluorescent Protein in Aot Reversed Micelles*. *Biophys Chem* **87**: 73-84.
- Vagin, A. y Teplyakov, A. (2010). *Molecular Replacement with Molrep*. *Acta Crystallogr D Biol Crystallogr* **66**: 22-25.
- Vaguine, A.A., Richelle, J. y Wodak, S.J. (1999). *Sfcheck: A Unified Set of Procedures for Evaluating the Quality of Macromolecular Structure-Factor Data and Their Agreement with the Atomic Model*. *Acta Crystallogr D Biol Crystallogr* **55**: 191-205.
- Vaillancourt, P., Zheng, C.F., Hoang, D.Q. y Breister, L. (2000). *Affinity Purification of Recombinant Proteins Fused to Calmodulin or to Calmodulin-Binding Peptides*. *Methods Enzymol* **326**: 340-362.
- Valencia, A. (2005). *Automatic Annotation of Protein Function*. *Curr Opin Struct Biol* **15**: 267-274.
- Valeur, B. y Berberan-Santos, M.N. (2012). *Molecular Fluorescence: Principles and Applications*, Wiley-VCH, Weinheim, Alemania

- Valeur, B. y Weber, G. (1977). *Resolution of the Fluorescence Excitation Spectrum of Indole into the 1la and 1lb Excitation Bands*. Photochem Photobiol **25**: 441-444.
- Van Den Berg, T.K., Honing, H., Franke, N., Van Remoortere, A., Schiphorst, W.E., Liu, F.T., Deelder, A.M., Cummings, R.D., Hokke, C.H. y Van Die, I. (2004). *Lacdinac-Glycans Constitute a Parasite Pattern for Galectin-3-Mediated Immune Recognition*. J Immunol **173**: 1902-1907.
- Van Den Eijnden, D.H., Neeleman, A.P., Van Der Knaap, W.P., Bakker, H., Agterberg, M. y Van Die, I. (1995). *Novel Glycosylation Routes for Glycoproteins: The Lacdinac Pathway*. Biochem Soc Trans **23**: 175-179.
- Van Kooyk, Y. y Rabinovich, G.A. (2008). *Protein-Glycan Interactions in the Control of Innate and Adaptive Immune Responses*. Nat Immunol **9**: 593-601.
- Varki, A. (1993). *Biological Roles of Oligosaccharides: All of the Theories Are Correct*. Glycobiology **3**: 97-130.
- Varrot, A., Basheer, S.M. y Imberty, A. (2013). *Fungal Lectins: Structure, Function and Potential Applications*. Curr Opin Struct Biol.
- Vasta, G.R. (2009). *Roles of Galectins in Infection*. Nat Rev Microbiol **7**: 424-438.
- Velázquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S. y Freire, E. (2004). *Isothermal Titration Calorimetry*. Curr Protoc Cell Biol Capítulo 17: Unidades 17 y 18.
- Villeneuve, N., Le Tacon, F. y Bouchard, D. (1991). *Survival of Inoculated Laccaria Bicolor in Competition with Native Ectomycorrhizal Fungi and Effects on the Growth of Outplanted Douglas-Fir Seedlings*. Plant and Soil **135**: 95-107.
- Voss, S. y Skerra, A. (1997). *Mutagenesis of a Flexible Loop in Streptavidin Leads to Higher Affinity for the Strep-Tag II Peptide and Improved Performance in Recombinant Protein Purification*. Protein Eng **10**: 975-982.
- Vrielink, A. y Sampson, N. (2003). *Sub-Angstrom Resolution Enzyme X-Ray Structures: Is Seeing Believing?* Curr Opin Struct Biol **13**: 709-715.
- Vyas, M.N., Vyas, N.K. y Quirocho, F.A. (1994). *Crystallographic Analysis of the Epimeric and Anomeric Specificity of the Periplasmic Transport/Chemosensory Protein Receptor for D-Glucose and D-Galactose*. Biochemistry **33**: 4762-4768.
- Walser, P.J., Kues, U., Aebi, M. y Kunzler, M. (2005). *Ligand Interactions of the Coprinopsis Cinerea Galectins*. Fungal Genet Biol **42**: 293-305.
- Wang, B. y Qiu, Y.L. (2006). *Phylogenetic Distribution and Evolution of Mycorrhizas in Land Plants*. Mycorrhiza **16**: 299-363.
- Wang, D., Liu, S., Trummer, B.J., Deng, C. y Wang, A. (2002). *Carbohydrate Microarrays for the Recognition of Cross-Reactive Molecular Markers of Microbes and Host Cells*. Nat Biotechnol **20**: 275-281.

- Wang, H. y Ng, T.B. (2005). *First Report of an Arabinose-Specific Fungal Lectin*. Biochem Biophys Res Commun **337**: 621-625.
- Wang, M., Trigueros, V., Paquereau, L., Chavant, L. y Fournier, D. (2002). *Proteins as Active Compounds Involved in Insecticidal Activity of Mushroom Fruitbodies*. J Econ Entomol **95**: 603-607.
- Ward, H.D. (1997). Glycobiology of Parasites: Role of Carbohydrate-Binding Proteins and Their Ligands in the Host-Parasite Interactions. *Glycosciences: Status and Perspectives*. H. J. Gabius and S. Gabius. Weinheim, Chapman & Hall: 399-413.
- Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S. y Tanaka, H. (1994). *The Roles of the C-Terminal Domain and Type Iii Domains of Chitinase A1 from Bacillus Circulans W1-12 in Chitin Degradation*. J Bacteriol **176**: 4465-4472.
- Watkins, W.M. (1999). *A Half Century of Blood-Group Antigen Research: Some Personal Recollections*. Trends Glycosci Glycotechnol **11**: 391-412.
- Watkins, W.M. (2001). *The Abo Blood Group System: Historical Background*. Transfus Med **11**: 243-265.
- Watson, K.A., Mitchell, E.P., Johnson, L.N., Son, J.C., Bichard, C.J., Orchard, M.G., Fleet, G.W., Oikonomakos, N.G., Leonidas, D.D., Kontou, M. y Et Al. (1994). *Design of Inhibitors of Glycogen Phosphorylase: A Study of Alpha- and Beta-C-Glucosides and 1-Thio-Beta-D-Glucose Compounds*. Biochemistry **33**: 5745-5758.
- Waugh, D.S. (2005). *Making the Most of Affinity Tags*. Trends Biotechnol **23**: 316-320.
- Weber, G. (1960). *Fluorescence-Polarization Spectrum and Electronic-Energy Transfer in Tyrosine, Tryptophan and Related Compounds*. Biochem J **75**: 335-345.
- Weis, W., Brown, J.H., Cusack, S., Paulson, J.C., Skehel, J.J. y Wiley, D.C. (1988). *Structure of the Influenza Virus Haemagglutinin Complexed with Its Receptor, Sialic Acid*. Nature **333**: 426-431.
- Weis, W.I. y Drickamer, K. (1996). *Structural Basis of Lectin-Carbohydrate Recognition*. Annu Rev Biochem **65**: 441-473.
- Wenzel, M., Gers-Barlag, H., Schimpl, A. y Rudiger, H. (1993). *Time Course of Lectin and Storage Protein Biosynthesis in Developing Pea (Pisum Sativum) Seeds*. Biol Chem Hoppe Seyler **374**: 887-894.
- White, R.H. (2006). *The Difficult Road from Sequence to Function*. J Bacteriol **188**: 3431-3432.
- Wimmerova, M., Mitchell, E., Sanchez, J.F., Gautier, C. y Imberty, A. (2003). *Crystal Structure of Fungal Lectin: Six-Bladed Beta-Propeller Fold and Novel Fucose Recognition Mode for Aleuria Aurantia Lectin*. J Biol Chem **278**: 27059-27067.



- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. y Wilson, K.S. (2011). *Overview of the Ccp4 Suite and Current Developments*. Acta Crystallogr D Biol Crystallogr **67**: 235-242.
- Wiseman, T., Williston, S., Brandts, J.F. y Lin, L.N. (1989). *Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter*. Anal Biochem **179**: 131-137.
- Wohlschlager, T., Butschi, A., Zurfluh, K., Vonesch, S.C., Auf Dem Keller, U., Gehrig, P., Bleuler-Martinez, S., Hengartner, M.O., Aebi, M. y Kunzler, M. (2011). *Nematotoxicity of Marasmius Oreades Agglutinin (Moa) Depends on Glycolipid Binding and Cysteine Protease Activity*. J Biol Chem **286**: 30337-30343.
- Woody, R.W. (1994). *Contributions of Tryptophan Side Chains to the Far-Ultraviolet Circular Dichroism of Proteins*. Eur Biophys J **23**: 253-262.
- Xu, Z., Bae, W., Mulchandani, A., Mehra, R.K. y Chen, W. (2002). *Heavy Metal Removal by Novel Cbd-Ec20 Sorbents Immobilized on Cellulose*. Biomacromolecules **3**: 462-465.
- Xu, Z., Horwich, A.L. y Sigler, P.B. (1997). *The Crystal Structure of the Asymmetric Groel-Groes-(Adp)7 Chaperonin Complex*. Nature **388**: 741-750.
- Yang, N., Li, D.F., Feng, L., Xiang, Y., Liu, W., Sun, H. y Wang, D.C. (2009). *Structural Basis for the Tumor Cell Apoptosis-Inducing Activity of an Antitumor Lectin from the Edible Mushroom Agrocybe Aegerita*. J Mol Biol **387**: 694-705.
- Yang, N., Tong, X., Xiang, Y., Zhang, Y., Sun, H. y Wang, D.C. (2005). *Crystallization and Preliminary Crystallographic Studies of the Recombinant Antitumour Lectin from the Edible Mushroom Agrocybe Aegerita*. Biochim Biophys Acta **1751**: 209-212.
- Yao, J., Nellas, R.B., Glover, M.M. y Shen, T. (2011). *Stability and Sugar Recognition Ability of Ricin-Like Carbohydrate Binding Domains*. Biochemistry **50**: 4097-4104.
- Young, N.M., Watson, D.C., Yaguchi, M., Adar, R., Arango, R., Rodriguez-Arango, E., Sharon, N., Blay, P.K. y Thibault, P. (1995). *C-Terminal Post-Translational Proteolysis of Plant Lectins and Their Recombinant Forms Expressed in Escherichia Coli. Characterization of "Ragged Ends" by Mass Spectrometry*. J Biol Chem **270**: 2563-2570.
- Zeilstra-Ryalls, J., Fayet, O. y Georgopoulos, C. (1991). *The Universally Conserved Groe (Hsp60) Chaperonins*. Annu Rev Microbiol **45**: 301-325.
- Zentz, C., Frenoy, J.P. y Bourrillon, R. (1978). *Binding of Galactose and Lactose to Ricin. Equilibrium Studies*. Biochim Biophys Acta **536**: 18-26.
- Zhang, F., Lees, E., Amin, F., Rivera Gil, P., Yang, F., Mulvaney, P. y Parak, W.J. (2011). *Polymer-Coated Nanoparticles: A Universal Tool for Biolabelling Experiments*. Small **7**: 3113-3127.

Zhang, Y.B., Howitt, J., Mccorkle, S., Lawrence, P., Springer, K. y Freimuth, P. (2004). *Protein Aggregation During Overexpression Limited by Peptide Extensions with Large Net Negative Charge*. *Protein Expr Purif* **36**: 207-216.

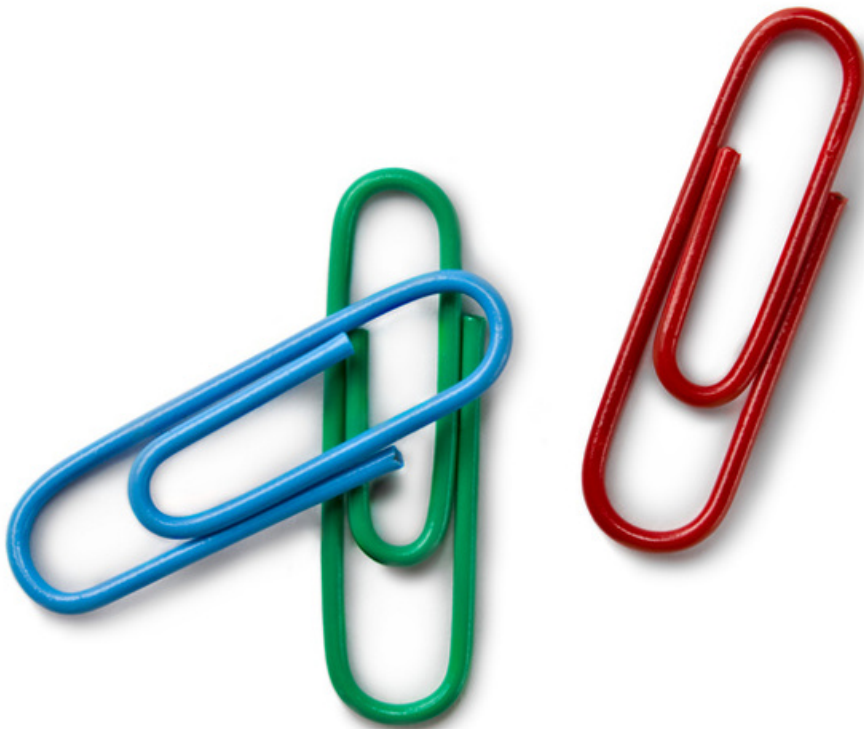
Zhao, J.K., Wang, H.X. y Ng, T.B. (2009). *Purification and Characterization of a Novel Lectin from the Toxic Wild Mushroom *Inocybe Umbrinella**. *Toxicon* **53**: 360-366.

Zhao, S., Guo, Y.X., Liu, Q.H., Wang, H.X. y Ng, T.B. (2009). *Lectins but Not Antifungal Proteins Exhibit Anti-Nematode Activity*. *Environ Toxicol Pharmacol* **28**: 265-268.

Zhuang, C., Murata, T., Usui, T., Kawagishi, H. y Kobayashi, K. (1996). *Purification and Characterization of a Lectin from the Toxic Mushroom *Amanita Pantherina**. *Biochim Biophys Acta* **1291**: 40-44.



# ANEXOS





## ANEXOS I-III

### Anexo I: Glosario

#### 1. Agarosa

Biopolímero, componente predominante del agar, constituido por la repetición de agarobiosa. La agarobiosa es un disacárido formado por unidades de D-galactosa y 3,6-anhidro-L-galactopiranosas unidas mediante enlaces  $\alpha(1\rightarrow3)$  y  $\beta(1\rightarrow4)$ .

#### 2. Aglutinación

Fenómeno de interacción cruzada entre células debido a la participación de un agente aglutinante (por ejemplo, una lectina) que conduce a la aglomeración desordenada de las células. Cuando ocurre entre glóbulos rojos, este proceso se conoce con el nombre de hemaglutinación.

#### 3. Anisotropía de fluorescencia

Fenómeno físico que se produce cuando la radiación emitida por un fluoróforo posee diferentes intensidades en diferentes ejes de polarización. La medida de la anisotropía de fluorescencia aporta información sobre la movilidad de la molécula y, por tanto, es útil para estudiar procesos dinámicos, como por ejemplo cambios conformacionales en proteínas.

#### 4. Anómero

Forma isomérica o epímero de un carbohidrato cíclico cuya diferencia radica única y exclusivamente en la configuración alrededor del átomo de carbono hemiacetalico. El átomo de carbono hemiacetalico se denomina carbono anomérico y el grupo hidroxilo unido a él puede presentar dos configuraciones:  $\alpha$  (posición axial) y  $\beta$  (posición ecuatorial).

## 5. Ascomycota

Hongos con micelio septado y dicariótico que producen ascosporas endógenas. Existen hongos ascomicetos con reproducción sexual o asexual y los más evolucionados presentan cuerpos fructíferos (ascocarpos).

## 6. Basidiomycota

Hongos superiores con micelio septado, principalmente dicariótico, que producen basidiosporas en los cuerpos fructíferos (basidiocarpos). Los hongos basidiomicetos son los más desarrollados e incluyen a las setas.

## 7. Carbohidrato

Azúcar, glúcido. Los carbohidratos son polihidroxialdehídos o polihidroxicetonas, o bien sustancias cuya hidrólisis da lugar a estos compuestos. La mayoría de ellos poseen la fórmula empírica  $(CH_2O)_n$ , aunque algunos contienen nitrógeno, fósforo o azufre. Los carbohidratos más sencillos se denominan monosacáridos, pero la repetición de éstos da lugar a la formación de cadenas más complejas, conocidas como glicanos. A su vez, éstos se dividen en oligosacáridos (de dos a diez repeticiones) y polisacáridos (con más de diez repeticiones por molécula).

## 8. Cristalografía de rayos X

Ciencia interdisciplinar que estudia las estructuras cristalinas, abarcando desde los procesos de formación de los cristales hasta la resolución estructural a nivel atómico de los mismos. La cristalografía de rayos X se fundamenta en el análisis de los patrones de difracción de los rayos X al atravesar una estructura cristalina. Actualmente, se considera una de las herramientas más potentes en campos como la Biología Estructural y la Ciencia de Materiales.

## 9. Cuerpos fructíferos

Estructura multicelular típica de los hongos, donde éstos llevan a cabo la producción y el almacenamiento de esporas. Los cuerpos fructíferos, además, participan activamente en el ciclo de reproducción sexual de los hongos. Existe un rango significativo de formas y morfologías, siendo los más característicos las setas de los hongos basidiomicetos.

## 10. Cuerpos proteicos

Sistema de almacenamiento de proteínas que actúan principalmente como fuente de nutrientes para las plantas durante procesos fisiológicos comprometidos como el desarrollo o la germinación del polen.

## 11. Epímero

Forma isomérica adoptada por una molécula que le permite distinguirse de otra por la diferente configuración alrededor de uno de sus centros estereogénicos o quirales.

## 12. Etiqueta de fusión

Péptido o proteína que se fusiona a una proteína de interés en el extremo N o C-terminal con el fin de optimizar la expresión recombinante y/o el método de purificación. Además, las etiquetas de fusión pueden mejorar la estabilidad, la resistencia a la degradación por proteasas y aumentar la solubilidad.

## 13. Factor *B*

Los factores *B* hacen referencia a los parámetros de desplazamiento atómico, es decir, a la medida del vector desplazamiento de un átomo con respecto a su posición media. Estos desplazamientos pueden deberse a diferentes efectos: vibración térmica, desorden estático en las diferentes moléculas de la unidad asimétrica, o efectos dinámicos del propio grupo al que pertenece el átomo. Los factores *B* pueden determinarse desde una perspectiva isótropa o anisótropa.

## 14. Fluoróforo

Fluorocromo. Molécula o parte de una molécula que posee fluorescencia. El fluoróforo es capaz de absorber energía de una determinada longitud de onda y volver a emitirla, pero con una longitud de onda mayor (menor energía). La cantidad de energía emitida y la longitud de onda dependerán tanto del propio fluoróforo como de su ambiente químico.

## 15. Furanosa

Término utilizado para designar la estructura química que adoptan algunos carbohidratos cuando en la reacción de ciclación dan lugar a la formación de un anillo de cinco eslabones, siendo uno de ellos un átomo de oxígeno y el resto átomos de carbono.



## 16. Galectina

Familia de lectinas caracterizada por presentar especificidad por  $\beta$ -galáctosidos. Las galectinas están extremadamente conservadas entre sí y todas ellas poseen, al menos, un dominio de unión de carbohidratos de ~140 aminoácidos.

## 17. Glicoconjugado

Término general para designar a un compuesto que ha sido modificado covalentemente con la incorporación de un carbohidrato, por ejemplo, una glicoproteína o un glicolípido.

## 18. Glicoma

Conjunto de carbohidratos o azúcares de un organismo, bien en su estado libre o bien formando parte de estructuras más complejas, como glicoproteínas o glicolípidos. Es una entidad compleja y variable que depende de multitud de factores, entre ellos, el genoma y el entorno.

## 19. Inmovilización de proteínas

La inmovilización es el proceso por el cual se lleva a cabo el confinamiento de proteínas (generalmente, enzimas) en una matriz sólida con el fin de aprovechar la actividad de la molécula para algún proceso de interés y poder usarla en repetidas ocasiones de forma continuada. La inmovilización suele ir ligada a un aumento de la estabilidad de la proteína y simplifica las etapas de purificación del producto (biocatálisis), aunque la complejidad del sistema aumenta considerablemente (problemas difusionales y orientación de la macromolécula).

## 20. Interacción C-H/ $\pi$

Fuerza intermolecular débil que tiene lugar entre los grupos C-H polarizados positivamente y la nube electrónica  $\pi$  de los sistemas aromáticos. El reconocimiento de los azúcares por parte de las proteínas generalmente implica el establecimiento de este tipo de interacciones entre residuos aromáticos del sitio de unión y parches hidrofóbicos del ligando.

**21. Interacción catión- $\pi$** 

Interacción molecular no covalente que se establece entre la cara de un sistema  $\pi$  rico en electrones (por ejemplo, un anillo aromático) con un catión adyacente (un metal o una molécula cargada). Esta interacción es tan intensa como un enlace de hidrógeno y juega un papel importante en el reconocimiento molecular.

**22. Lectina**

Proteína caracterizada por la capacidad que presenta para unir carbohidratos de forma específica. Las lectinas están ampliamente distribuidas en la naturaleza y presentan gran diversidad estructural. Este grupo de proteínas excluye a anticuerpos, enzimas que actúan sobre carbohidratos o proteínas transportadoras de mono y/u oligosacáridos. Suelen tener actividad hemaglutinante y participan generalmente en procesos de reconocimiento celular, señalización y adhesión. En la actualidad son consideradas una herramienta muy útil en los laboratorios de Biología Molecular y representan una fuerte promesa para el desarrollo de aplicaciones biotecnológicas de interés.

**23. Lectina de tipo C**

Grupo de lectinas caracterizado por requerir ion calcio ( $\text{Ca}^{2+}$ ) para el reconocimiento del ligando. Las lectinas de tipo C abarcan gran variedad de funciones biológicas.

**24. Ligando multivalente**

Un ligando multivalente es aquel ligando que puede ser reconocido simultáneamente por diferentes sitios de unión de una molécula de proteína (multivalencia) o por varias proteínas al mismo tiempo.

**25. Micorriza**

Simbiosis entre un hongo y las raíces de una planta, donde ambos participantes obtienen beneficios, principalmente nutrientes.

**26. Multivalencia**

Capacidad que poseen algunas lectinas para reconocer ligandos mediante dos o más sitios de unión por molécula de proteína bien por la presencia de más de un sitio de unión por molécula o bien por la capacidad asociativa para formar multímeros funcionales.

## **27. Microarray**

Matriz constituida por una colección de lectinas (*lectin microarray*) o de azúcares (*glycan microarray*) inmovilizados sobre una superficie sólida en un orden determinado que puede utilizarse para analizar las interacciones lectina-azúcar y determinar así o la naturaleza química de un ligando o la especificidad de una lectina, respectivamente.

## **28. Nanopartícula magnética**

Tipo de nanopartícula que puede ser manipulada mediante el uso de campos magnéticos. Generalmente, este tipo de partículas está constituido por un núcleo magnético de hierro, níquel, oro o cobalto derivatizado con diferentes compuestos químicos que les otorgan funcionalidad. Suelen tener un tamaño comprendido entre 5-500 nm y sus propiedades dependen generalmente del método de síntesis y la estructura química que adoptan.

## **29. Nemátodo**

Gusano con aparato digestivo en forma de conducto recto que ocupa toda la extensión del cuerpo. La mayoría son organismos acuáticos y viven de forma autónoma o como parásitos de otros seres vivos.

## **30. Piranosa**

Término utilizado para designar la estructura química que adoptan algunos carbohidratos cuando forman un anillo de seis eslabones, siendo uno de ellos un átomo de oxígeno y el resto átomos de carbono.

## **31. Proteína de fusión**

Proteína creada mediante ingeniería genética a partir de la fusión de dos o más genes que originariamente codificaban proteínas independientes. El uso de proteínas de fusión es una herramienta frecuentemente empleada en la producción de proteínas recombinantes, especialmente porque la unión de etiquetas de fusión facilita la expresión y/o la purificación de las mismas.

## **32. Pseudosimetría**

Relaciones geométricas que aparecen en algunas proteínas resultado de la similitud estructural entre subdominios y/o protómeros, que no necesariamente comparten una evidente similitud de secuencia.

### **33. Radiación sincrotrón**

Radiación electromagnética de alta energía emitida en los anillos de las instalaciones de tipo sincrotrón. Los electrones (o positrones), que viajan a grandes velocidades en canales rectilíneos, son obligados a cambiar de dirección mediante el uso de campos magnéticos de gran energía. En ese momento, los electrones emiten radiación muy energética y policromática. El uso de monocromadores y el ajuste de la longitud de onda permite que en las instalaciones de tipo sincrotrón se puedan hacer ensayos de dispersión anómala (SAD y MAD).

### **34. Reemplazo molecular**

Método de determinación estructural empleado para resolver el problema de la fase que utiliza una combinación de los factores de estructura experimentales con las fases de un modelo, cuya identidad de secuencia es suficientemente elevada (~30% o más) para asegurar que ambas estructuras van a ser homólogas.

### **35. Subsitio**

Los subsitios o sitios extendidos en las lectinas son sitios de reconocimiento que rodean al sitio principal de unión del ligando e interaccionan con éste, estableciendo enlaces con más de un residuo del azúcar, lo que incrementa la afinidad por los oligo y polisacáridos.

### **36. Transferencia de energía resonante de Förster**

La transferencia de energía resonante de Förster (FRET) es un mecanismo de transferencia de energía entre cromóforos. La excitación de un cromóforo puede dar lugar a la emisión de energía de otro cromóforo cercano si existe un acoplamiento dipolo-dipolo que implica el solapamiento del espectro de emisión del primero con el de excitación del segundo.

### **37. Trébol $\beta$**

Plegamiento adoptado por algunas proteínas, principalmente lectinas, que consiste en la repetición de tres subdominios ( $\alpha$ ,  $\beta$  y  $\gamma$ ), constituido cada uno de ellos por una lámina de cuatro hebras  $\beta$  antiparalelas. En su conjunto, forman un barril  $\beta$  (dominio  $\lambda$ ) coronado

por un triplete de horquillas  $\beta$  que posee pseudosimetría de orden 3 a lo largo del eje principal.

## Anexo II: Abreviaturas

ADN	Ácido desoxirribonucleico
ADNc	ADN codificante
ADP	Parámetros de desplazamiento anisótropo ( <i>Anisotropic Displacement Parameters</i> )
ARN	Ácido ribonucleico
ARNm	ARN mensajero
ANL	Lectina de <i>Aspergillus niger</i>
BEL	Lectina de <i>Boletus edulis</i>
Bgl	$\beta$ -galactosidasa de <i>Lactobacillus plantarum</i>
BTL2	Lipasa 2 de <i>Geobacillus thermocatenolatus</i>
CBD <sup>1</sup>	Dominios de unión a celulosa ( <i>Cellulose-Binding Domains</i> )
CBD <sup>2</sup>	Dominios de unión a quitina ( <i>Chitin-Binding Domains</i> )
CCL2	Isolectina 2 de <i>Coprinopsis cinérea</i>
CFG	Consorcio de Glicómica Funcional ( <i>Consortium for Functional Glycomics</i> )
CLSM	Microscopía confocal láser de barrido ( <i>Confocal Laser Scanning Microscopy</i> )
CNL	Lectina de <i>Clitocybe nebularis</i>
Cpl-7	Endolisina del bacteriófago Cp-7
CRD	Dominio de unión a carbohidratos ( <i>Carbohydrate Recognition Domain</i> )
DAB	3,3'-diaminobenzidina
DC	Dicroísmo circular
DTT	1,4-ditiotreitol
DO <sub>600</sub>	Densidad óptica a 600 nm
DSC	Calorimetría diferencial de barrido ( <i>Differential Scanning Calorimetry</i> )
ECM	Ectomicorriza
EDTA	Ácido etilendiaminotetraacético

EGFP	Proteína verde fluorescente mejorada ( <i>Enhanced Green Fluorescent Protein</i> )
ESRF	Instalación europea de radiación sincrotrón ( <i>European Synchrotron Radiation Facility</i> )
FB	Cuerpo fructífero ( <i>Fruiting Body</i> )
FBL	Lectinas de los cuerpos fructíferos ( <i>Fruiting Body Lectins</i> )
FPLC	Cromatografía rápida de proteínas en fase líquida ( <i>Fast Protein Liquid Chromatography</i> )
FRET	Transferencia de energía de resonancia de Förster ( <i>Förster Resonance Energy Transfer</i> )
FTR	Motivos repetidos por partida triple en estructuras de tipo trébol $\beta$ ( $\beta$ - <i>Trefoil Four Three-fold Repetitive Motifs</i> )
Gal	Galactosa
GBP	Proteína de unión a azúcares ( <i>Glycan Binding Protein</i> )
GGG	$\beta$ -3'-galactosillactosa
Glc	Glucosa
GlcNAc	N-acetilglucosamina
GST	Glutación S-transferasa
HEPES	Ácido 4-(2-hidroxietil)-1-piperazinetanosulfónico
HRP	Peroxidasa de rábano ( <i>Horseradish peroxidase</i> )
IPTG	Isopropiltiogalactósido
IRM	Imagen por resonancia magnética
ITC	Calorimetría de valoración isotérmica ( <i>Isothermal Titration Calorimetry</i> )
Lac	Lactosa
LacNAc	N-acetillactosamina
LBL	Lectina de <i>Laccaria bicolor</i>
LBL <sub>152</sub>	Módulo lectina N-terminal de la lectina de <i>Laccaria bicolor</i>
LIC	Clonación independiente de ligasa ( <i>Ligase Independent Cloning</i> )
LSL	Lectina de <i>Laetiporus sulphureus</i>
LSLa	Isolectina A de <i>Laetiporus sulphureus</i>
LSLb	Isolectina B de <i>Laetiporus sulphureus</i>
LSLc	Isolectina C de <i>Laetiporus sulphureus</i>

LSLt	Etiqueta de fusión basada en el módulo lectina N-terminal de LSLa (LSL <sub>150</sub> + <i>linker</i> + sitio de reconocimiento TEV)
LSL <sub>150</sub>	Módulo lectina N-terminal de la isolectina A de <i>Laetiporus sulphureus</i>
MBP	Proteína de unión a maltosa ( <i>Maltose-Binding Protein</i> )
MOA	Aglutinina de <i>Marasmius oreades</i>
MPD	2-metil-2,4-pentanodiol
NCS	Simetría no cristalográfica ( <i>Non-Crystallographic Symmetry</i> )
Neu5Ac	Ácido siálico
NPA	Nanopartículas magnéticas recubiertas de agarosa
ORF	Marco de lectura abierta ( <i>Open Reading Frame</i> )
PAGE	Electroforesis en geles de poliacrilamida ( <i>Polyacrylamide Gel Electrophoresis</i> )
PCR	Reacción en cadena de la polimerasa ( <i>Polimerase Chain Reaction</i> )
PFM	Módulo C-terminal formador de poros de la isolectina A de <i>Laetiporus sulphureus</i> ( <i>Pore-Forming Module</i> )
PSL	Lectina de <i>Polyporus squamosus</i>
RFU	Unidades de fluorescencia relativa ( <i>Relative Fluorescence Units</i> )
RIP	Proteínas inactivadoras de ribosomas ( <i>Ribosome-Inactivating Proteins</i> )
RMN	Resonancia magnética nuclear
RMSD	Desviación cuadrática media ( <i>Root-Mean-Square Deviation</i> )
RSA	Aglutinina de <i>Rhizoctonia solani</i>
rpm	Revoluciones por minuto
SDS	Dodecilsulfato sódico
SSA	Aglutinina de <i>Sclerotinia sclerotiorum</i>
SSL	Lectina de <i>Sclerotinia sclerotiorum</i>
TEM	Microscopía electrónica de transmisión ( <i>Transmission Electron Microscopy</i> )
TEV	Virus del grabado del tabaco ( <i>Tobacco Etch Virus</i> )
TLS	Translación-Libración-Desplazamiento helicoidal ( <i>Translation-Libration-Screw</i> )
Tris	2-amino-2-hidroximetilpropano-1,3-diol
UA	Unidad asimétrica
UV	Ultravioleta



WS	Sitios de mayor probabilidad para albergar moléculas de agua ( <i>Water Sites</i> )
----	---

## Anexo III: Publicaciones

### Publicaciones derivadas del presente trabajo

López-Gallego, F.; **Acebrón, I.**; Mancheño, J.M.; Raja, S.; Lillo, M.P. y Guisán Seijas J.M. (2012) *Directed, strong, and reversible immobilization of proteins tagged with a  $\beta$ -trefoil lectin domain: a simple method to immobilize biomolecules on plain agarose matrixes*. Bioconjug. Chem. 23: 565-73.

Angulo, I.; **Acebrón, I.**; de las Rivas, B.; Muñoz, R.; Rodríguez-Crespo, I.; Menéndez, M.; García, P.; Tateno, H.; Goldstein, I.J.; Pérez-Agote, B. y Mancheño, J.M. (2011) *High-resolution structural insights on the sugar-recognition and fusion tag properties of a versatile  $\beta$ -trefoil lectin domain from the mushroom *Laetiporus sulphureus**. Glycobiology. 21: 1.349-61.

**Acebrón, I.**; Curiel, J.A.; de las Rivas, B.; Muñoz, R. y Mancheño, J.M. (2009) *Cloning, production, purification and preliminary crystallographic analysis of a glycosidase from the food lactic acid bacterium *Lactobacillus plantarum* CECT 748(T)*. Protein Expr. Purif. 68: 177-82.

